



**HAL**  
open science

# Commande $H_\infty$ paramétrique et application aux viseurs gyrostabilisés

Guillaume Rance

► **To cite this version:**

Guillaume Rance. Commande  $H_\infty$  paramétrique et application aux viseurs gyrostabilisés. Automatique / Robotique. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS152 . tel-01904086

**HAL Id: tel-01904086**

**<https://theses.hal.science/tel-01904086v1>**

Submitted on 24 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Commande $H_\infty$ paramétrique et application aux viseurs gyro-stabilisés

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

Ecole doctorale n°580  
Sciences et Technologies de l'Information et de la Communication (STIC)  
Spécialité de doctorat : Automatique

Thèse présentée et soutenue à CentraleSupélec, Gif-sur-Yvette, le 9 Juillet 2018, par

**M. GUILLAUME RANCE**

Composition du Jury :

M. Gilles Duc Professeur, CentraleSupélec (L2S)	Président du jury
M. Olivier Bachelier Professeur, Université de Poitiers (LIAS)	Rapporteur
M. François Boulier Professeur, Université de Lille (CRISTAL)	Rapporteur
M. Jean-Jacques Loiseau Directeur de recherches, Université de Nantes (LS2N)	Rapporteur
Mme. Mioara Joldes Chargée de recherches, CNRS (LAAS)	Examinatrice
M. Guillaume Moroz Chargé de recherches, Inria Nancy - Grand Est	Examineur
M. Alban Quadrat Directeur de recherches, Inria Lille - Nord Europe	Directeur de thèse
M. Arnaud Quadrat Ingénieur, Safran Electronics & Defense	Encadrant de thèse
M. Hugues Mounier Professeur, Université Paris-Saclay (L2S)	Invité
M. Serge Hirwa Ingénieur, Safran Electronics & Defense	Invité



# Remerciements

L'opportunité d'effectuer cette thèse est née d'un stage de fin d'études réalisée aux côtés d'Arnaud Quadrat. Durant ce stage, j'ai pu découvrir le monde de la recherche avec une personne passionnée, qui a su me transmettre ce goût pour la science et qui m'a donné l'envie de continuer dans cette voie. Cette passion, je l'ai retrouvée chez son frère, Alban Quadrat. Ces deux frères m'ont appris énormément de choses durant cette thèse, autant d'un point de vue scientifique, que philosophique ou humain. Merci mille fois pour cet encadrement exceptionnel que vous m'avez accordé durant ces années. Ce fut une expérience inoubliable, et c'est grâce à vous.

J'aimerais remercier infiniment mon maître Jedi du calcul formel, Yacine Bouzidi, qui aura su prendre le temps de m'expliquer encore et encore toutes ces notions qui sont au cœur de cette thèse. Tu as joué un grand rôle pendant ces trois années, et a su me remonter le moral dans des moments de doute. Aujourd'hui, je suis content de ne pas te considérer comme un simple collègue, mais comme un très bon ami. Merci pour tout.

Mes pensées vont ensuite naturellement aux nombreux membres de ce jury. J'aimerais tout d'abord remercier Olivier Bachelier d'avoir accepté d'être rapporteur. J'espère que ce mémoire t'aura appris des choses sur la commande  $H_\infty$ , toi qui fut le premier à me l'enseigner. Merci à Jean-Jacques Loiseau et François Boulter de m'accorder leur temps pour noter ce travail. Merci à Mioara Joldes, Gilles Duc, Guillaume Moroz et Hugues Mounier de compter parmi les examinateurs de ce jury pluridisciplinaire. Un remerciement particulier à Serge Hirwa, dernier membre de ce jury, pour sa patience et son intérêt pour mon travail, ainsi que pour ces premiers moments enrichissants passés ensemble sur projet.

J'aimerais ensuite remercier toutes les personnes m'ayant entouré pendant ces trois années. Merci au pôle systèmes asservis d'avoir su m'intégrer dans l'équipe. Stéphane pour ton côté rock and roll! Émilie pour ton côté potins! Marine, pour ta bonne humeur quotidienne, une qualité incroyable... Et, bien évidemment, merci aussi à Pierre, Franck, Philippe, Lambert, Maxime, Virginie pour leur sympathie.

Merci aussi à toutes les personnes que j'ai croisé à Supélec, avec qui j'ai pu partager des moments et qui ont donné de belles amitiés. Remerciements particuliers à Walid et Dina, meilleurs co-bureaux de Supélec, pour les fous rires durant les longues après midi d'hiver. Merci à Lucien pour son soutien, tu es et resteras un vrai bro! Je te souhaite beaucoup de réussite pour ta carrière, tu le mérites! Merci à Hugo de m'avoir invité à ce poker, de là est né une belle amitié! Merci à François pour ta présence durant toute cette thèse. Merci à Lucie d'avoir partagé des rires tonitruants. Enfin, merci aussi à Maria, Yoan, Sophie, Gordana, Caetano, ... La liste est trop longue pour tous vous citer!

Je voudrais aussi remercier mes amis rencontrés lors de ces moments dansants. Vous avez réussi à me changer les idées lors de ces moments difficiles. Un grand merci donc à Adrien, Ele, Alice, Sarah, Delphine! J'espère vous recroiser à l'avenir pour une danse!

Merci aussi à ceux qui continuent à garder le contact malgré la distance. Jean Phi, merci pour tes nombreuses cartes postales! Mais surtout, merci pour tes qualités humaines incroyables, je te souhaite que du bonheur. Thomich', toi qui est toujours là pour mettre un kudos sur strava! Merci énormément, je te souhaite toute la réussite que tu mérites.

---

Enfin, j'aimerais terminer par remercier ma famille sans qui rien n'aurait été possible. Merci Maman et Papa de m'avoir donné cette chance de réaliser ce que j'avais envie de faire. Vous m'avez toujours soutenu dans mes choix, et m'avez donné sans compter. Sans vous, rien de tout ceci n'aurait été possible. Je vous aime.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>15</b>
1.1	Contexte . . . . .	15
1.2	Qu'est-ce qu'un viseur ? . . . . .	16
1.3	Boucle de stabilisation inertielle . . . . .	18
1.4	Modélisation des viseurs . . . . .	20
1.5	Choix d'une méthode de synthèse de loi de commande . . . . .	22
1.6	État de l'art sur la résolution d'équations de Riccati algébriques et la commande $H_\infty$ paramétrique . . . . .	24
1.7	Plan de thèse . . . . .	26
1.8	Liste de publications . . . . .	27
<b>2</b>	<b>Fondements de la commande <math>H_\infty</math> par Loop-Shaping</b>	<b>29</b>
2.1	Introduction . . . . .	30
2.2	Caractérisation fréquentielle des objectifs de performance et de robustesse . . . . .	32
2.2.1	Fonction de sensibilité et contraintes de performance . . . . .	32
2.2.2	Fonction de sensibilité complémentaire et contraintes de robustesse . . . . .	32
2.2.3	Fonction de régulation et contraintes de performance en régulation . . . . .	34
2.2.4	Fonction de commande et contraintes de saturation . . . . .	34
2.3	Stabilisation de systèmes aux facteurs copremiers incertains via critère $H_\infty$ . . . . .	35
2.3.1	Factorisations copremières d'une matrice de transfert . . . . .	35
2.3.2	Un problème de commande robuste . . . . .	36
2.3.3	Conception d'un correcteur . . . . .	38
2.3.4	Cas particulier des systèmes SISO . . . . .	40
2.4	Lien avec la commande Linéaire Quadratique Gaussienne . . . . .	43
2.5	Commande $H_\infty$ par Loop-Shaping . . . . .	44
2.5.1	Modélage de la boucle ouverte . . . . .	44
2.5.2	Une procédure de synthèse de correcteur $H_\infty$ par Loop-Shaping . . . . .	47
2.6	Exemple d'application . . . . .	48
2.7	Conclusion . . . . .	52
<b>3</b>	<b>Approche symbolique des équations de Riccati algébriques</b>	<b>53</b>
3.1	Quelques notions d'algèbre . . . . .	55
3.1.1	Polynômes univariés et multivariés . . . . .	55
3.1.2	Extension algébrique et clôture algébrique . . . . .	55
3.1.3	Systèmes polynomiaux . . . . .	56
3.1.4	Classes d'équivalence . . . . .	58
3.1.5	Ordres monomiaux . . . . .	59
3.1.6	Division de polynômes à plusieurs variables . . . . .	60
3.1.7	Bases de Gröbner . . . . .	61

3.1.8	Matrices de multiplication . . . . .	63
3.1.9	Résultants . . . . .	64
3.1.10	Variétés discriminantes . . . . .	65
3.2	Classe de systèmes et équation de Riccati algébrique $\mathcal{R} = 0$ étudiés . . . . .	69
3.3	Réduction de $\mathcal{R} = 0$ à un système d'équations polynomiales $\mathcal{B} = 0$ . . . . .	70
3.3.1	Paramétrisation des solutions de $\mathcal{R} = 0$ par un calcul des sous-espaces propres de la matrice Hamiltonienne associée . . . . .	70
	Propriétés de la matrice Hamiltonienne associée à l'EAR . . . . .	70
	Calcul d'une solution de $\mathcal{R} = 0$ à paramètres fixés . . . . .	79
	Calcul d'une solution de $\mathcal{R} = 0$ à paramètres non-fixés par factorisation spectrale . . . . .	80
3.3.2	Paramétrisation des solutions de $\mathcal{R} = 0$ par une étude directe du système d'équation polynomiales . . . . .	88
3.3.3	Équivalence des deux approches . . . . .	96
3.4	Propriétés du système polynomial $\mathcal{E}$ issu de $\mathcal{R} = 0$ . . . . .	96
3.5	Étude de $\mathcal{E}$ à paramètres fixés . . . . .	98
3.5.1	Exemple d'introduction . . . . .	98
3.5.2	Éléments séparant $V(\langle \mathcal{B} \rangle)$ . . . . .	99
3.5.3	Paramétrisation de $V(\langle \mathcal{B} \rangle)$ . . . . .	103
3.5.4	Caractérisation de $X > 0$ solution de $\mathcal{R} = 0$ . . . . .	106
3.5.5	Isolation certifiée des zéros de $\mathcal{B}$ . . . . .	109
3.6	Étude de $\mathcal{E}$ à paramètres non fixés . . . . .	110
3.6.1	Remarques préliminaires . . . . .	110
3.6.2	Paramétrisation des solutions de $\mathcal{E}$ pour $1 \leq n \leq 4$ . . . . .	111
3.6.3	Solutions de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$ . . . . .	113
3.6.4	Solution définie positive de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$ . . . . .	113
3.7	Cas particulier des systèmes comportant une approximation de Padé . . . . .	122
3.7.1	Rappels sur les approximations de Padé . . . . .	122
3.7.2	Approximations de Padé et factorisation spectrale . . . . .	123
3.8	Conclusion . . . . .	124
<b>4</b>	<b>Approche symbolique de la commande <math>H_\infty</math> par Loop-Shaping</b> . . . . .	<b>127</b>
4.1	Résolution explicite du Problème de Commande Robuste . . . . .	128
4.1.1	Calcul explicite du critère $\gamma_{\text{opt}}$ . . . . .	128
4.1.2	Correcteurs $H_\infty$ paramétriques . . . . .	138
4.1.3	Exemples . . . . .	140
	Système d'ordre 2 : masse-ressort avec extrémité encastrée . . . . .	140
	Système d'ordre 3 : modèle double-poutre d'un viseur gyrostabilité en 2D . . . . .	143
	Système d'ordre 4 : système masse-ressort libre . . . . .	146
4.2	Résolution symbolique/numérique du problème de commande robuste . . . . .	152
4.3	Réglage optimal d'une pondération Loop-Shaping . . . . .	155
4.3.1	Une méthode de réglage de pondération scalaire de type gain . . . . .	156
4.3.2	Application au système double masse-ressort . . . . .	159
	Étude du système sans amortissement : réglage de pondération global . . . . .	159
	Étude du système avec amortissement : réglage local de pondération . . . . .	165
4.4	Commande $H_\infty$ adaptative indirecte . . . . .	170
4.4.1	Système double masse-ressort sans amortissement bouclé avec correction adaptative globale . . . . .	171

4.4.2	Système double masse-ressort avec amortissement bouclé avec correction adaptative locale . . . . .	171
4.5	Conclusion . . . . .	173
<b>5</b>	<b>Extension de la commande <math>H_\infty</math> paramétrique aux systèmes à retard</b>	<b>175</b>
5.1	Décomposition additive de systèmes à retard et stabilisation . . . . .	176
5.2	Problème de commande robuste et décomposition additive . . . . .	178
5.2.1	Liens entre les projecteurs $\Pi(K, G)$ et $\Pi(K_0, G_0)$ . . . . .	178
5.2.2	Un encadrement de $\ \Pi(K, G)\ _\infty$ . . . . .	180
5.2.3	Application à un modèle de viseur gyrostabilisé . . . . .	182
5.3	Réglage d'un correcteur robuste sous-optimal à l'aide d'une pondération Loop-Shaping . . . . .	189
5.3.1	Système réel et hypothèses . . . . .	190
5.3.2	Système fictif et données préliminaires . . . . .	191
5.3.3	Une procédure de conception de correcteur Loop-Shaping adaptatif stabilisant le $G$ . . . . .	192
5.3.4	Caractéristiques de l'asservissement obtenu . . . . .	197
5.4	Conclusion . . . . .	204
<b>6</b>	<b>Conclusion et perspectives</b>	<b>205</b>
6.1	Résumé des contributions . . . . .	205
6.2	Perspectives . . . . .	207
<b>A</b>	<b>Systèmes linéaires continus invariants dans le temps</b>	<b>209</b>
A.1	Définition d'un système linéaire . . . . .	209
A.2	Stabilité d'un système linéaire . . . . .	210
A.3	Commandabilité et observabilité . . . . .	210
A.4	Stabilisabilité et détectabilité . . . . .	211
A.5	Formes compagnes et matrices de passage . . . . .	212
<b>B</b>	<b>Normes <math>H_2</math> et <math>H_\infty</math></b>	<b>215</b>
B.1	Norme euclidienne . . . . .	215
B.2	Normes matricielles . . . . .	215
B.3	Normes de signaux . . . . .	215
B.4	Normes de systèmes . . . . .	216
B.4.1	Espace $H_2(\mathbb{C}_+)$ et norme $H_2$ . . . . .	216
B.4.2	Espace $H_\infty(\mathbb{C}_+)$ et norme $H_\infty$ . . . . .	217
	Calcul numérique . . . . .	217
	Calcul symbolique . . . . .	218
	Inégalités classiques . . . . .	218
<b>C</b>	<b>Matrices définies positives</b>	<b>219</b>
C.1	Produits de matrices définies positives . . . . .	219
<b>D</b>	<b>Racines par radicaux de polynômes de degré inférieur ou égal à 4</b>	<b>221</b>
D.1	Polynômes de degré 2 . . . . .	221
D.1.1	Racines par radicaux . . . . .	221
D.1.2	Racine réelle maximale . . . . .	221
D.2	Polynômes de degré 3 . . . . .	222
D.2.1	Racines par radicaux . . . . .	223

• Cas $\mu_1 = 0$ .	223
• Cas $\mu_1 \neq 0$ .	223
Récapitulatif	225
D.2.2 Racine réelle maximale	226
Avec des radicaux	226
Avec fonctions trigonométriques	228
D.3 Polynômes de degré 4	229
D.3.1 Racines par radicaux	229
• Cas $\mu_1 = 0$ .	229
• Cas $\mu_1 \neq 0$ .	229
Récapitulatif	231
D.3.2 Racine réelle maximale	231
<b>E Complexité algorithmique et binaire d'une RUR de <math>\mathcal{E}</math></b>	<b>235</b>
E.1 Définitions et notations	235
E.2 Résultant multivarié, $u$ -résultant et RUR	236
E.3 Complexité algorithmique et binaire d'une RUR	237
E.4 Application au système $\mathcal{E}$	238
<b>F Modélisation d'un viseur en 2 dimensions</b>	<b>239</b>
F.1 Hypothèses et notations	239
F.2 Équations de la mécanique	240
F.3 Flexibilité	242
F.3.1 Flexibilité exacte	242
F.3.2 Cas idéal de flexibilité pour deux ressorts égaux	243
F.3.3 Cas de deux ressorts égaux sans amortissement	244
<b>G Identification de fonction de transfert</b>	<b>245</b>
G.1 Mesures hors ligne à disposition	245
G.2 Identification du système	246
G.2.1 Une première modélisation simple avec retard	247
G.2.2 Une modélisation d'ordre 4 sans retard	248
G.2.3 Une modélisation d'ordre 5 plus fine	249
G.2.4 Modélisations plus fines	249
G.3 Récapitulatif des différentes modélisations	250

# Table des figures

1.1	Viseurs Safran Electronics & Defense – Copyright Safran . . . . .	16
1.2	Différentes applications des viseurs – Copyright Safran . . . . .	16
1.3	Viseur “miror stabilized” . . . . .	17
1.4	Viseur “mass stabilized” . . . . .	17
1.5	Boucle de stabilisation inertielle de Ligne de Visée . . . . .	18
1.6	Boucle de stabilisation . . . . .	19
1.7	Modèle double-poutre d’un viseur gyrostabilisé . . . . .	20
1.8	Réponse fréquentielle de $G$ . . . . .	22
2.1	Représentation d’un système bouclé . . . . .	30
2.2	Boucle de stabilisation . . . . .	31
2.3	Incertitudes de modélisation . . . . .	33
2.4	Stabilisation d’une famille de modèles $\mathcal{G}_\varepsilon$ . . . . .	36
2.5	Construction du correcteur central : équivalence de deux écritures . . . . .	39
2.6	Modelage de la boucle ouverte . . . . .	47
2.7	Diagramme de Bode des échantillons de modèles $P$ choisis dans $\mathcal{B}$ . . . . .	49
2.8	Diagramme de Bode de la pondération Loop-Shaping utilisée . . . . .	50
2.9	Diagramme de Bode du correcteur $C$ obtenu . . . . .	50
2.10	Prise en compte d’incertitudes de modèle via la sensibilité complémentaire . . . . .	51
2.11	Diagramme de Black de $K G$ pour tout $G \in \mathcal{B}$ . . . . .	51
3.1	Variété discriminante de $P$ suivant $b_1$ . . . . .	67
3.2	Schéma du spectre de $\mathcal{H}$ avec $\lambda \in \mathbb{C}_+$ vérifiant $f(\lambda) = 0$ . . . . .	79
3.3	Système masse-ressort encastré . . . . .	85
3.4	Schéma de substitutions des éléments de $X$ . . . . .	91
3.5	Modèle deux-masses-ressort . . . . .	95
3.6	Représentation de $V(\mathcal{B})$ dans le cas $n = 2$ , $d_0 = 1/4$ , $d_2 = 2$ . . . . .	98
3.7	Représentation de $V(\mathcal{B})$ dans le cas $n = 2$ , $d_0 = 1/4$ , $d_2 = 2$ (croix rouges), des directions de projection des formes linéaires non séparant $V(\mathcal{B})$ (droites jaunes) et de la direction de projection de la forme séparante $t : (b_0, b_1) \mapsto b_0 - 2b_1$ (droite bleue) . . . . .	100
4.1	$\gamma_{\text{opt}}(\nu)$ associé au système $G = \frac{c_0}{s+a_0}$ , où $\nu = \frac{a_0}{c_0}$ . . . . .	133
4.2	Système masse-ressort encastré . . . . .	140
4.3	$\gamma_{\text{opt}}$ du système masse-ressort encastré . . . . .	141
4.4	Diagramme de Bode de $K_{\gamma_{\text{opt}}}$ pour $m = 1$ et $k = 1$ . . . . .	142
4.5	Diagramme de Black-Nichols de $K_{\gamma_{\text{opt}}} G$ pour $m = 1$ et $k = 1$ . . . . .	142
4.6	Modèle double-poutre d’un viseur gyrostabilisé . . . . .	143
4.7	$\gamma_{\text{opt}}$ pour un système double poutre . . . . .	144
4.8	Diagramme de Bode de $K_{\gamma_{\text{opt}}}$ pour $J_1 = J_2 = 1$ , $k = 1$ et $l = 1$ . . . . .	145

4.9	Diagramme de Black-Nichols de $K_{\gamma_{\text{opt}}} G$ pour $J_1 = J_2 = 1$ , $k = 1$ et $l = 1$ . . .	145
4.10	Modèle deux-masses-ressort . . . . .	146
4.11	Tracé de $\gamma_{\text{opt}}$ en fonction de $a_2$ et $c_0$ . . . . .	149
4.12	Tracé de $\gamma_{\text{opt}}$ en fonction de $G_r$ et $\rho$ . . . . .	151
4.13	Système deux masses ressort amorti . . . . .	159
4.14	Décomposition de l'espace des paramètres à l'aide de la variété discriminante de $L_t$ . . . . .	163
4.15	$L(\lambda, G_r, \rho) = 0$ pour $\rho \in \{0, 1/100, 1/10\}$ . . . . .	167
4.16	Bleu + vert : $\bar{L}(G_r, \rho) = 0$ - Vert : $\gamma_{\text{opt}}(G_r, \rho) = \bar{\gamma}_{\text{opt}}$ - Rouge : $G_r = \hat{G}_r(\rho)$ .	168
4.17	Bleu + vert : $L(\gamma^2 - 1, \hat{G}_r(\rho), \rho) = 0$ - Vert : branche correspondant à $\gamma_{\text{opt}}(\hat{G}_r(\rho), \rho)$ - Rouge : $\gamma_{\text{opt}}$ attendu . . . . .	169
4.18	Commande adaptative indirecte . . . . .	170
4.19	Simulation du système sans amortissement bouclé soumis à une variation paramétrique . . . . .	172
4.20	Simulation du système avec amortissement bouclé soumis à une variation paramétrique . . . . .	172
5.1	Schéma d'asservissement d'un système avec retard . . . . .	177
5.2	Schéma d'asservissement avec $\Delta = 0$ . . . . .	177
5.3	Évolution de $\ D\ _{\infty}^2$ et $\ D\ _{\infty}^{-2}$ en fonction de $\ \Delta\ _{\infty}$ . . . . .	182
5.4	Décomposition en cellules de l'espace des paramètres $\{\xi, r\}$ . . . . .	186
5.5	Variations de $ \bar{R}(i\Omega) $ pour $\xi < \sqrt{2}/2$ . . . . .	187
5.6	$ \bar{R}(i\Omega) $ - Rouge : $\xi = 0.1$ , $r = 1.01$ - Bleue : $\xi = 0.1$ , $r = 0.99$ . . . . .	187
5.7	Diagramme de Bode de $R$ et sa norme $H_{\infty}$ . . . . .	188
5.8	Diagramme de Bode du système $P$ pour $\omega_1 \in \Omega_{\star}$ . . . . .	190
5.9	Tracé de $\gamma_{0_{\text{opt}}}(\theta)$ . . . . .	192
5.10	Tracé du diagramme de Black-Nichols de $PC$ pour $w = 82.82$ et $\omega_1 = 2$ . . .	193
5.11	Bleue : $\gamma_{0_{\text{opt}}}(g, \omega_1) = 1.7$ - Rouge : $g = \hat{g}(\omega_1)$ . . . . .	194
5.12	Graphe de $\gamma_{0_{\text{opt}}}(\hat{g}(\omega_1), \omega_1)$ . . . . .	195
5.13	Graphe de $\hat{w}(\omega_1)$ . . . . .	195
5.14	Procédure de Loop-Shaping pour un système à retard . . . . .	196
5.15	Correcteur $C(s, \omega_1)$ obtenu pour différentes valeurs de $\omega_1$ . . . . .	199
5.16	Diagramme de Bode de la boucle ouverte $C(\omega_1)P(\omega_1)$ obtenu pour $\omega_1 \in \Omega_{\star}$ .	199
5.17	Diagramme de Black-Nichols de $C(\omega_1)P(\omega_1)$ obtenu pour $\omega_1 \in \Omega_{\star}$ . . . . .	200
5.18	Diagramme de Black-Nichols de $C(\omega_1)P(\omega_1)$ obtenu pour différentes valeurs de $\omega_1$ lorsque $C$ est fixé au correcteur nominal. Bleu : $\omega_1 = 1.1$ (instable) - Violet : $\omega_1 = 2$ (stable) - Rouge : $\omega_1 = 3$ (instable) . . . . .	200
5.19	Diagramme de Bode de $S(\omega_1) = \frac{1}{1+C(\omega_1)P(\omega_1)}$ obtenu pour $\omega_1 \in \Omega_{\star}$ . . . . .	201
5.20	Diagramme de Bode de $P(\omega_1)S(\omega_1) = \frac{P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$ obtenu pour $\omega_1 \in \Omega_{\star}$ . .	201
5.21	Diagramme de Bode de $C(\omega_1)S(\omega_1) = \frac{C(\omega_1)}{1+C(\omega_1)P(\omega_1)}$ obtenu pour $\omega_1 \in \Omega_{\star}$ . .	202
5.22	Diagramme de Bode de $T(\omega_1) = \frac{C(\omega_1)P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$ obtenu pour $\omega_1 \in \Omega_{\star}$ . . . . .	202
5.23	Réponse à un échelon de $\frac{P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$ obtenue pour $\omega_1 \in \Omega_{\star}$ . . . . .	203
5.24	Réponse à un échelon de $\frac{C(\omega_1)P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$ obtenue pour $\omega_1 \in \Omega_{\star}$ . . . . .	203
B.1	Norme $H_{\infty}$ d'une fonction de transfert $G$ dans le cas SISO . . . . .	217
D.1	Variété discriminante $\delta_2 = 0$ de $P_2$ . . . . .	222

F.1	Modélisation d'un viseur gyrostabilisé par une double poutre . . . . .	239
F.2	Schéma simplifié d'un viseur en 2D . . . . .	240
G.1	Système de stabilisation <i>Bloc Gyro Miroir (BGM)</i> . . . . .	245
G.2	Réponse fréquentielle de $y/u = G_m P H$ . . . . .	246
G.3	Réponse fréquentielle de $y/i = K_m P H$ . . . . .	247
G.4	Modélisation avec une inertie, un mode mécanique et un retard . . . . .	248
G.5	Modélisation avec une inertie, un mode mécanique et une approximation de Padé d'ordre 1 . . . . .	249
G.6	Modélisation avec une inertie, un mode mécanique, un passe-bas et un retard pur ou défini par une approximation de Padé d'ordre 1 . . . . .	250
G.7	Modélisation avec une inertie, un mode mécanique, un passe-bas et un retard pur ou défini par une approximation de Padé d'ordre 1 . . . . .	251
G.8	Récapitulatif des différentes modélisations . . . . .	251



# Liste des tableaux

3.1	Matrices $X$ solution de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$ . . . . .	118
3.2	Système polynomial $\mathcal{B}$ pour $1 \leq n \leq 4$ . . . . .	118
3.3	Expression des $d_{2k}$ en fonction de $a$ et $c$ pour $1 \leq n \leq 4$ . . . . .	119
3.4	Paramétrisation rationnelle $\tilde{\mathcal{R}}$ des solutions de $\tilde{\mathcal{E}}$ pour $1 \leq n \leq 4$ . . . . .	119
3.5	Racines par radicaux de $\tilde{p}$ pour $1 \leq n \leq 4$ . . . . .	120
3.6	Variétés singulières introduites par les racines par radicaux de $\tilde{p}$ . . . . .	121
3.7	Expression par radicaux de la racine réelle maximale de $\tilde{p}$ pour $2 \leq n \leq 4$ . . . . .	121
4.1	Conditions de non-observabilité $\Delta_o = 0$ pour $1 \leq n \leq 4$ . . . . .	133
4.2	Matrices $\mathcal{Q}$ pour $1 \leq n \leq 4$ . . . . .	134
4.3	Matrices $Y$ pour $1 \leq n \leq 4$ . . . . .	135
4.4	Polynôme caractéristique $h$ de la matrice $YX$ pour $1 \leq n \leq 4$ . . . . .	136
4.5	Formes closes de $\gamma_{\text{opt}}$ pour $1 \leq n \leq 4$ . . . . .	137
4.6	$\gamma_{\text{opt}}$ , marges de phase $m_\Phi$ et de gain $m_G$ garanties (bornes inférieures), et marges de phase $M_\Phi$ et de gain $M_G$ réelles pour $\xi \in \{0, 1/2, 1\}$ . . . . .	141
4.7	$\gamma_{\text{opt}}$ , marges de phase $m_\Phi$ et de gain $m_G$ garanties, et marges de phase $M_\Phi$ et de gain $M_G$ réelles pour $\xi \in \{0, 1/2, 1\}$ . . . . .	144
D.1	Test . . . . .	226
G.1	Modélisations possibles du système mesuré . . . . .	252

## Abréviations

BIBO	Bounded Input Bounded Output, c'est-à-dire entrée bornée sortie bornée
ERA	Équation de Riccati Algébrique
LC	Leading Coefficient, c'est-à-dire coefficient de tête
LM	Leading Monomial, c'est-à-dire monôme de tête
LMI	Linear Matrix Inequality, c'est-à-dire Inégalité Matricielle Linéaire
LT	Leading Term, c'est-à-dire terme de tête
LTI	Linéaire à Temps Invariant
LQ	Linéaire Quadratique
LQG	Linéaire Quadratique Gaussienne
MIMO	Multiple Input Multiple Output, c'est-à-dire multi entrées multi sorties
PCR	Problème de Commande Robuste
SISO	Single Input Single Output, c'est-à-dire simple entrée simple sortie

## Notations

$\mathbb{N}$	Ensemble des entiers naturels : $\{0, 1, 2, \dots\}$
$\mathbb{Z}$	Ensemble des entiers relatifs
$\llbracket a, b \rrbracket$	Ensemble des entiers compris entre $a$ et $b$
$\mathbb{Q}$	Ensemble des nombres rationnels
$\mathbb{R}$	Ensemble des nombres réels
$\mathbb{C}$	Ensemble des nombres complexes
$\mathbb{C}_+$	Ensemble des nombres complexes à partie réelle strictement positive
$\mathbb{C}_-$	Ensemble des nombres complexes à partie réelle strictement négative
$\mathbb{K}_*$	Ensemble $\mathbb{K}$ privé de l'élément nul, $\mathbb{K}$ désignant $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$
$\mathbb{K}_+$	Ensemble des éléments positifs ou nuls de $\mathbb{K}$ , $\mathbb{K}$ désignant $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$
$\mathbb{K}_-$	Ensemble des éléments négatifs ou nuls de $\mathbb{K}$ , $\mathbb{K}$ désignant $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$
$\mathbb{K}[x_1, \dots, x_n]$	Ensemble des polynômes en $x_1, \dots, x_n$ et à coefficients dans $\mathbb{K}$
$\mathbb{K}(x_1, \dots, x_n)$	Ensemble des fractions rationnelles en $x_1, \dots, x_n$ et à coefficients dans $\mathbb{K}$
$i$	$\sqrt{-1}$
$j$	$-\frac{1}{2} + i \frac{\sqrt{3}}{2}$ - Nombre complexe vérifiant $1 + j + j^2 = 1$
$x^{1/n}$	Désigne la racine $n^{\text{ème}}$ principale de $x \in \mathbb{C}$ , où $n \in \mathbb{Q}$ , c'est-à-dire $\exp\left(\frac{1}{n} (\ln( x ) + i \arg(x))\right)$ où $\arg(x)$ est l'argument principal de $x$ , c'est-à-dire tel que $-\pi \leq \arg(x) \leq \pi$
$\lambda_{\max}(M)$	Valeur propre réelle maximale d'une matrice $M$
$\lambda_*(M)$	Maximum de l'ensemble des valeurs absolues des valeurs propres réelles d'une matrice $M$
$s$	Variable de Laplace
$\text{Tr}(M)$	Trace de la matrice $M$
$M^T$	Transposée de la matrice $M$
$M^*$	Transposée conjuguée de $M \in \mathbb{R}^{m \times p}$ , c'est-à-dire $M^*(s) = M^T(-s)$
$\ G\ _\infty$	Norme $H_\infty$ de la matrice de transfert $G$
$\ M\ _H$	Norme de Hankel de la matrice de transfert $M$
$H_\infty$	$H_\infty(\mathbb{C}_+)$ : algèbre des fonctions holomorphes dans $\mathbb{C}_+$ bornées pour la norme $\ f\ _\infty = \sup_{s \in \mathbb{C}_+}  f(s) $
$RH_\infty$	Espace des fonctions rationnelles propres et stables

# Chapitre 1

## Introduction

### Sommaire

---

1.1	Contexte . . . . .	15
1.2	Qu'est-ce qu'un viseur ? . . . . .	16
1.3	Boucle de stabilisation inertielle . . . . .	18
1.4	Modélisation des viseurs . . . . .	20
1.5	Choix d'une méthode de synthèse de loi de commande . . . . .	22
1.6	État de l'art sur la résolution d'équations de Riccati algébriques et la commande $H_\infty$ paramétrique . . . . .	24
1.7	Plan de thèse . . . . .	26
1.8	Liste de publications . . . . .	27

---

### 1.1 Contexte

Cette thèse s'inscrit au sein d'une collaboration entre *Safran Electronics & Defense* (ex. *Sagem Défense Sécurité*), filiale du groupe *Safran*, l' *Institut National de Recherche en Informatique et Automatique* et le *Laboratoire des Signaux et Systèmes*.

L'objectif de cette thèse est de développer des techniques nouvelles d'asservissement de viseurs via des méthodes récentes issues de la recherche académique. Cette thèse vient compléter les travaux de plusieurs thèses développés à *Safran Electronics & Defense* en collaboration avec le milieu académique :

- la thèse de P. Coustal [Cou97] est la première thèse proposant d'appliquer les techniques  $H_\infty$  aux viseurs ;
- la thèse de S. Hirwa [Hir13] propose une méthode de synthèse de correcteurs robustes à ordre fixe ;
- les travaux de P. Feyel [Fey15] développent des techniques d'optimisation de correcteurs par les métaheuristiques ;
- la thèse de S. Frasnado [Fra16] étudie l'optimisation de lois de commande d'un imageur sur critère optronique, etc.

Afin de définir clairement le problème pratique que l'on cherche à résoudre dans cette thèse, il est tout d'abord nécessaire de décrire ce qu'est un *viseur* et sa *Ligne de Visée*. Nous exposerons ensuite différentes problématiques émergeant de ces technologies et des moyens existants pour les résoudre. Finalement, nous motiverons cette thèse et expliquerons la démarche adoptée dans ce mémoire.

## 1.2 Qu'est-ce qu'un viseur ?

Un *viseur* est un ensemble de caméras et/ou de dispositifs de pointage, appelé *cœur optique*, monté sur un porteur généralement en mouvement et orientable dans l'espace. On appelle alors *Ligne de Visée* (LdV) l'axe optique sortant de l'un de ces capteurs. Des viseurs sont représentés en Figures 1.1 et 1.2. Ces équipements permettent la recherche, la localisation, l'identification et la désignation de cible, le suivi de cible, de l'aide au pilotage, etc. Le rôle premier du viseur est de stabiliser et d'orienter la LdV quels que soient les mouvements du porteur et/ou de la cible, et quel que soit l'environnement extérieur (conditions atmosphériques, obstructions, etc.).

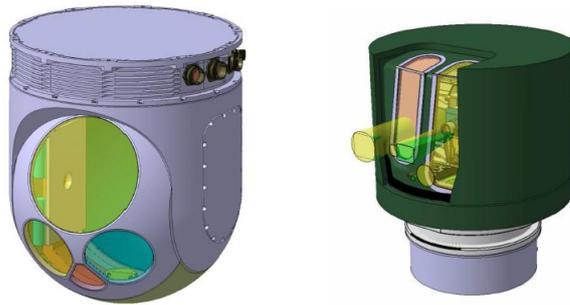


FIGURE 1.1 – Viseurs Safran Electronics & Defense – Copyright Safran



FIGURE 1.2 – Différentes applications des viseurs – Copyright Safran

L'objectif premier du viseur est d'observer les détails d'une scène lointaine sur un environnement vibrant. Tout d'abord, expliquons pourquoi il est nécessaire de stabiliser la LdV. Pour cela, faisons le parallèle avec une situation de tous les jours. Tout le monde a remarqué que lorsque l'on désire prendre une photographie de nuit, il ne faut pas bouger pendant la prise de vue afin d'obtenir une photographie nette. En effet, les photons venant du paysage sont envoyés sur une grille de photo-capteurs. Un point lumineux est alors représenté par une impulsion lumineuse. Par rapport à une scène fixe, le mouvement de l'appareil photo va créer un déplacement relatif de cette impulsion sur la grille pendant le temps d'intégration du capteur. L'impulsion initiale devient alors une tâche qui donnera une *image floue* si celle-ci s'étend sur plusieurs pixels. Dans le cas de l'appareil photo, il est plus fréquent d'obtenir une image floue de nuit que de jour car l'appareil va naturellement augmenter le temps d'intégration du capteur lorsque la luminosité est faible. Il se passe exactement la même chose dans le cadre des viseurs.

Quelques aspects diffèrent cependant entre la prise d'une photographie dans la vie de tous les jours et l'observation précise d'une scène lointaine avec un environnement vibrant. Safran travaille avec des caméras à très *haute résolution* et avec des très *petits champs de vision* afin de "voir" loin. De plus, le viseur est monté sur un *porteur fortement vibrant*, ce qui va déformer la structure mécanique du viseur. Tous ces aspects font que l'on doit mettre en place un processus permettant de *stabiliser l'image de manière précise*. Les mouvements angulaires de la LdV recherchés sont de l'ordre du *micro-radian* ( $10^{-6}$  radians).

Pour cela, l'idée est de mettre en place un asservissement capable d'agir sur la vitesse/position angulaire de la LdV. La mesure de la position/vitesse angulaire de la LdV se fait au moyen d'un gyromètre ou d'une centrale inertielle embarquée sur le porteur combinée aux capteurs angulaires de la plateforme. La correction de la vitesse/position angulaire de la LdV se fait au moyen de *cardans* ou de *miroirs* actionnés par des *moteurs*. On distingue alors deux architectures possibles : les viseurs "*mass-stabilized*" et "*mirror-stabilized*" [Hil08].

Les viseurs "*mirror-stabilized*" (voir Figure 1.3) stabilisent la LdV par l'intermédiaire de miroirs orientables selon un ou deux axes, tandis que le cœur optronique du viseur reste solidaire du porteur. L'asservissement est alors réalisé à l'aide d'une centrale de navigation placée sur le porteur et de capteurs angulaires mesurant les orientations des miroirs par rapport à leurs supports respectifs. L'avantage de ce type de viseur est qu'il est protégé et demande des moteurs plus petits. Pour des applications militaires, un autre avantage est que le cœur du viseur peut être placé dans une enceinte blindée, et seuls les miroirs sont exposés aux tirs extérieurs.

Les viseurs "*mass-stabilized*" (voir Figure 1.4) se basent sur une stabilisation directe du cœur optronique. Le cœur est alors monté sur des cardans qui permettent de l'orienter suivant deux angles afin de balayer tout l'espace. Un gyroscope/gyromètre, placé proche du cœur, fournit une image de la position/vitesse angulaire de la LdV. Cette mesure est alors utilisée pour asservir les moteurs entraînant les cardans et permettant la stabilisation de la LdV. Ces types de viseurs ont généralement de meilleures performances que les miroirs stabilisés car ils travaillent en rejet de perturbation et non en suivi de consigne, ce qui impose de plus faibles contraintes temporelles sur les synchronisations des mesures et des consignes.

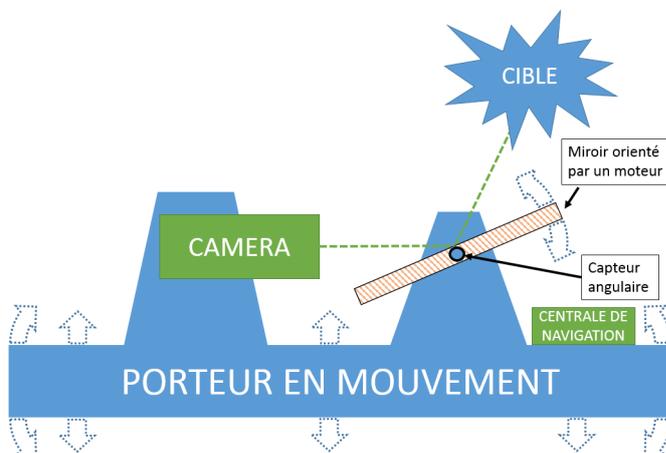


FIGURE 1.3 – Viseur "mirror stabilized"

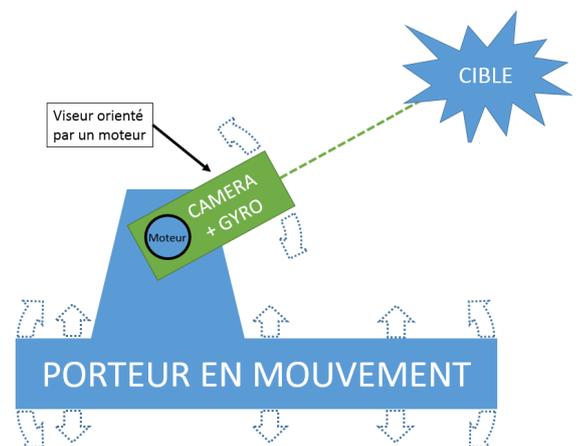


FIGURE 1.4 – Viseur "mass stabilized"

En pratique, certains viseurs comportent deux *étages de stabilisation* : pour chacune des rotations, un axe gros et un axe fin. Ceci permet de distinguer les fonctions de *stabilisation* et de *poursuite* : les axes gros permettent la poursuite de la cible (déplacement grossier de la LdV), tandis que l'axe fin permet la stabilisation de la ligne de visée (déplacement fin de la LdV).

Quelle que soit l'architecture choisie, il est nécessaire d'*asservir la LdV*, c'est-à-dire d'utiliser les mesures de position/vitesse angulaires fournies par les capteurs afin d'agir sur la position/vitesse angulaire de la LdV via les moteurs, et quelles que soient les perturbations extérieures. C'est sur cet aspect des viseurs que ce travail de thèse se focalise.

### 1.3 Boucle de stabilisation inertielle

Afin d'assurer une LdV *fixe dans un référentiel inertielle*, il est nécessaire de rejeter les perturbations agissant sur le cœur optronique. Pour cela, il faut que la somme des couples appliqués au cœur soit nulle. Ceci doit être le cas quels que soient les mouvements du porteur (angulaires ou vibrations linéaires) et quelles que soient les perturbations climatiques extérieures (variations de température, pression, ...). L'objectif final recherché est donc que la LdV soit fixe dans un repère inertielle, c'est-à-dire ait une vitesse angulaire nulle par rapport à ce repère considéré comme fixe. Nous parlerons alors improprement de *stabilisation inertielle* pour signifier que la LdV est fixe dans un repère inertielle (ce qui ne doit pas être confondu avec une notion de stabilité d'un asservissement).

Pour atteindre cet objectif, on met en place la *structure d'asservissement* décrite par la Figure 1.5. Cette boucle d'asservissement est composée d'*actionneurs* permettant d'agir sur un ensemble de pièces mécaniques à stabiliser, tel qu'un cœur optronique, dont on mesure la vitesse inertielle  $\dot{\theta}$  à l'aide d'un *capteur* (tel qu'un gyromètre).

Chaque bloc de la Figure 1.5 peut-être vu comme un *système*, c'est-à-dire un ensemble de relations liant des entrées et des sorties. On parlera aussi de *transfert* entre des entrées et des sorties. Par exemple, un bloc moteur peut être considéré comme un transfert entre une tension de commande  $u$  et un couple  $\Gamma_u$ . Pour plus de détails sur ces notions, nous nous référons à l'Annexe A.

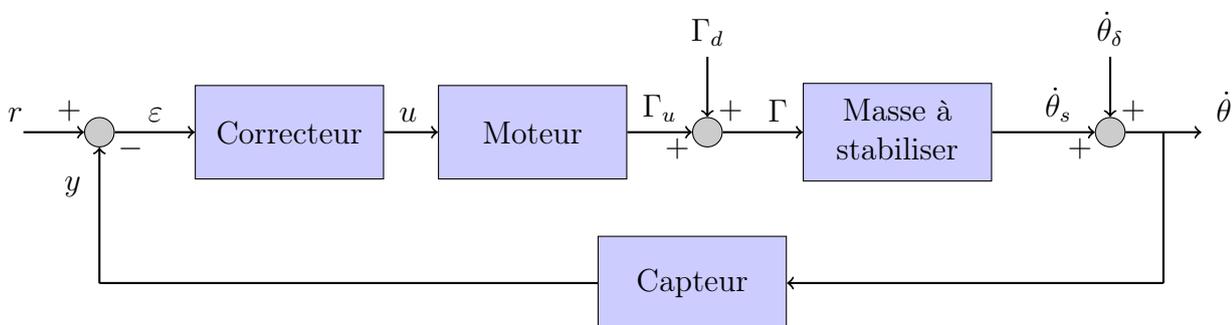


FIGURE 1.5 – Boucle de stabilisation inertielle de Ligne de Visée

Le but de la structure d'asservissement est de permettre aux moteurs de générer un couple  $\Gamma_u$  qui compense les couples perturbateurs  $\Gamma_d$ . La consigne  $u$  de ces moteurs est générée par

un correcteur. Cette consigne est une fonction de l'erreur d'asservissement  $\varepsilon$ , définie comme la différence entre une référence  $r$  (choisie nulle dans une problématique de stabilisation) et la mesure  $y$ . Plus cette erreur  $\varepsilon$  est importante, plus le couple  $\Gamma_u$  fourni par les moteurs devra être important afin de réduire cette erreur.

Savoir quel couple demander aux moteurs en fonction de l'erreur  $\varepsilon$  n'est pas trivial. Plusieurs phénomènes physiques viennent complexifier cette question. Premièrement, la masse à stabiliser n'est pas infiniment rigide et *résonne* à certaines fréquences d'excitation induites par les vibrations du porteur. De plus, comme nous le mentionnions précédemment, l'asservissement doit être *robuste vis-à-vis des perturbations extérieures*  $\Gamma_d$  et  $\dot{\theta}_\delta$  :  $\Gamma_d$  représente des *couples perturbateurs* (*disturbance* en anglais) induits par le porteur et  $\dot{\theta}_\delta$  symbolise des *déformations mécaniques* entre l'emplacement du capteur et l'emplacement des caméras à stabiliser. Par ailleurs, le capteur fournit une mesure  $y$  de la vitesse angulaire  $\dot{\theta}$  de la LdV avec un certain *retard* de transmission. La loi de commande doit donc être conçue pour prendre en compte ces différents phénomènes.

Mathématiquement parlant,  $\dot{\theta}_\delta$  peut être vue comme une perturbation agissant directement sur la mesure  $y$ . Nous pouvons alors regrouper la partie *Moteur*, *Masse à stabiliser* et *Capteur* en un unique système  $G$ , et considérer le schéma d'asservissement de la Figure 1.6. Celle-ci décrit simplement l'asservissement d'un système  $G$  soumis à des perturbations en entrée  $d$  et en sortie  $\delta$ , bouclé par un correcteur  $K$  ( $d$  étant reliée à  $\Gamma_d$  par l'inverse du transfert du moteur, et  $\delta$  à  $\dot{\theta}_\delta$  par l'inverse du transfert du capteur). C'est ce formalisme que l'on gardera tout au long de ce mémoire.

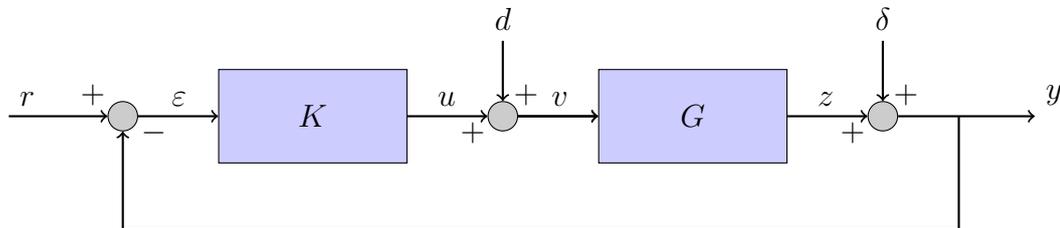


FIGURE 1.6 – Boucle de stabilisation

**Remarque 1.1.** Un viseur comporte deux axes de rotation. Dans la suite, nous ne prendrons pas en compte le couplage dynamique entre les deux axes, ni même le couplage induit par les déformées modales. Ainsi, avec ces approximations, l'asservissement d'un axe de rotation peut se faire de manière indépendante.

**Remarque 1.2.** Nous nous focaliserons sur le problème de la stabilisation inertielle de la LdV, c'est-à-dire nous ne détaillons que la boucle utilisée pour un asservissement en vitesse angulaire. Cependant, le cas de la *poursuite*, c'est-à-dire de l'asservissement en position angulaire, est (fortement) similaire à celui détaillé ici.

Il reste maintenant à concevoir une loi de commande des moteurs, c'est-à-dire un correcteur  $K$  qui satisfasse un cahier des charges. Bien qu'il existe des techniques permettant de concevoir un correcteur sans aucune connaissance du système (voir par exemple [FJ08] et les références citées), celles-ci restent limitées face à des techniques plus classiques se basant sur une connaissance a priori d'un *modèle* du système étudié. Pour cela, dans le prochain paragraphe, nous allons étudier la modélisation des viseurs.

## 1.4 Modélisation des viseurs

La *modélisation* d'un système physique peut se faire de deux manières : soit par la description de phénomènes physiques, soit par des mesures expérimentales. Dans le cadre des viseurs, nous disposons de différents moyens de modéliser ceux-ci.

Commençons par une modélisation simple en deux dimensions. Le choix de passer en deux dimensions se justifie en considérant que les deux axes de rotation d'un viseur sont (en première approximation) *découplés*. Sous une telle hypothèse, la rotation d'un axe n'influence pas la rotation de l'autre. C'est ce qui est considéré dans la majeure partie des études R&D à *Safran Electronics & Defense*.

Les liaisons pivot sont réalisées à partir de roulements à billes. De plus, les pièces mécaniques composant le cadre du viseur doivent être considérées comme déformables. On peut alors assimiler l'ensemble des liaisons et des pièces mécaniques comme des liaisons élastiques modélisées par des ressorts dans la direction radiale du roulement.

Ainsi, nous pouvons considérer la modélisation 2D de la Figure 1.7. Cette modélisation consiste en deux poutres indéformables reliées par deux ressorts de raideur et d'amortissement potentiellement différents. Les liaisons élastiques représentent les déformations de la structure mécanique du viseur. La première poutre représente le cadre du viseur entraîné en rotation par les moteurs. La deuxième poutre représente le cœur du viseur où sont placés les optiques et le gyromètre. La problématique de stabilisation consiste à stabiliser la deuxième poutre sachant que la première est soumise à un couple.

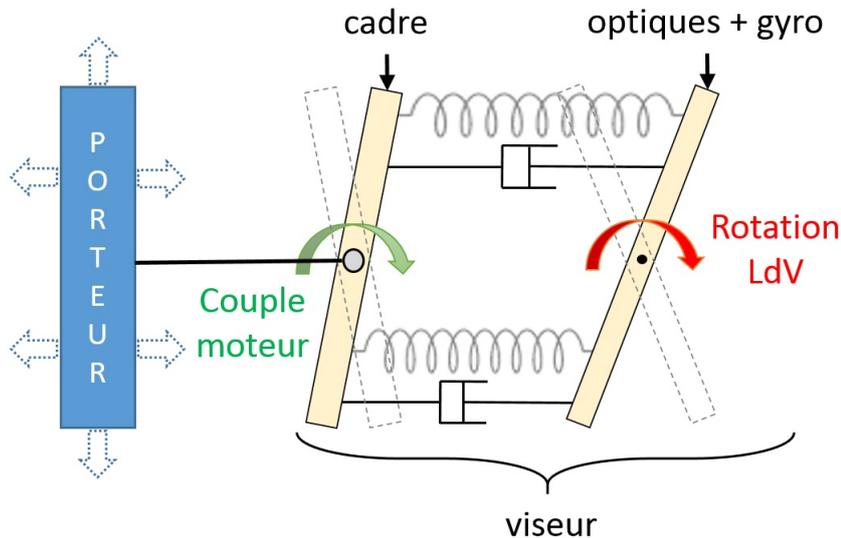


FIGURE 1.7 – Modèle double-poutre d'un viseur gyrostabilisé

Une telle représentation permet de comprendre quelques phénomènes physiques liés aux viseurs. Notamment, nous pouvons voir que si l'on entraîne en rotation la poutre de gauche, alors la poutre de droite n'effectue pas la même rotation à cause de la compression/extension

des ressorts. Le transfert entre le couple moteur et la rotation de la LdV est appelé *flexibilité* [GR03]. Par ailleurs, imaginons que la poutre de gauche soit entraînée en translation linéaire par le porteur (mouvement gauche droite). Alors, si les deux ressorts sont de raideurs différentes, la poutre de droite est entraînée en rotation. C'est ce que l'on appelle la *transmissibilité* [GR03].

La dynamique d'une telle modélisation peut se décrire par des *équations différentielles* (voir Annexe F). Nous pouvons en extraire des modèles sous forme de *fonction de transfert* dépendant de la variable de Laplace  $s \in \mathbb{C}$  (opérateur de dérivation). Par exemple, la fonction de transfert entre le couple mécanique exercé sur le cadre et la rotation de la Ligne de Visée a la forme suivante

$$G = \frac{c_0 + c_1 s + c_2 s^2 + c_3 s^3}{s^2 (a_2 + a_3 s + a_4 s^2 + a_5 s^3 + a_6 s^4)},$$

où les  $a_i$  et  $c_i$  sont des coefficients dépendant des raideurs et amortissements des ressorts, des inerties des poutres et de leurs longueurs (voir Corollaire F.1 de l'Annexe F). Lorsque les deux ressorts sont identiques, la fonction de transfert  $G$  se simplifie pour donner

$$G = \frac{n_0 (p^2 + 2\epsilon p s)}{s^2 (p^2 + 2\epsilon p s + s^2)},$$

où  $p$ ,  $\epsilon$  et  $n_0$  sont des grandeurs réelles. Dans cette modélisation,  $\epsilon$  est petit et nous voyons clairement apparaître un mode mécanique présentant une résonance et une inertie.

Les viseurs sont en général des structures complexes présentant plusieurs modes mécaniques. La modélisation présentée précédemment ne permet de prendre en compte qu'un seul mode mécanique. Elle est donc assez limitée.

Une alternative est de considérer des *modélisations par éléments finis* d'un viseur [GR03, Chap. 3.3]. Celles-ci consistent à représenter la structure mécanique par un ensemble de points, appelés nœuds, reliés par des ressorts. Plus le nombre de nœuds est important, plus la modélisation est complexe et modélise plus finement le système. Certains logiciels (tels que ANSYS [Bha02]) permettent d'extraire une représentation d'état du système à partir de l'étude de son comportement dynamique en s'appuyant sur une *analyse modale* du modèle [GR03]. Nous pouvons alors nous servir de cette représentation d'état pour concevoir un correcteur.

Enfin, lorsque l'on dispose de prototypes, il est possible de réaliser directement des mesures de fonctions de transfert du système. Ces mesures sont réalisées en utilisant un *transféromètre*. Le système linéaire à identifier  $G$  étant bouclé par un correcteur, le transféromètre permet de sommer une excitation sinusoïdale en entrée de système (que l'on peut considérer comme une perturbation en entrée de modèle) et de mesurer à la fois le couple moteur total appliqué au système ainsi que sa sortie. L'excitation sinusoïdale générée étant de fréquence connue  $f_0$ , les mesures sont aussi des sinus de même fréquence, et l'on peut estimer le gain et la phase de ceux-ci de manière à déduire  $G(i 2 \pi f_0) = |G(i 2 \pi f_0)| e^{i \arg(G(i 2 \pi f_0))}$ . Il suffit alors de réaliser des mesures pour différents  $f_0$  afin d'obtenir une série de points permettant d'approcher la fonction de transfert réelle du système. De ces séries de mesures, nous obtenons des courbes telles que celles données en Figure 1.8. Ces courbes nous permettent d'identifier plusieurs modèles valables sur une plage de fréquence plus ou moins grande (voir Annexe G). Par exemple, nous estimons que le transfert représenté en Figure 1.8 peut s'approcher (jusqu'à la fréquence du premier mode) par une fonction de transfert d'ordre 4 avec

une fréquence normalisée au numérateur et avec un retard :

$$G_{4,r} := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{1}{1 + \tau_1 s} \right) e^{-T_{4,r} s},$$

$$G_0 = 6.65 \cdot 10^{-3}, \quad T_{4,r} = 1.21, \quad \omega_0 = 1, \quad \omega_1 = 1.03, \quad \xi = 1.08 \cdot 10^{-2}, \quad \tau_1 = 1/1.14.$$

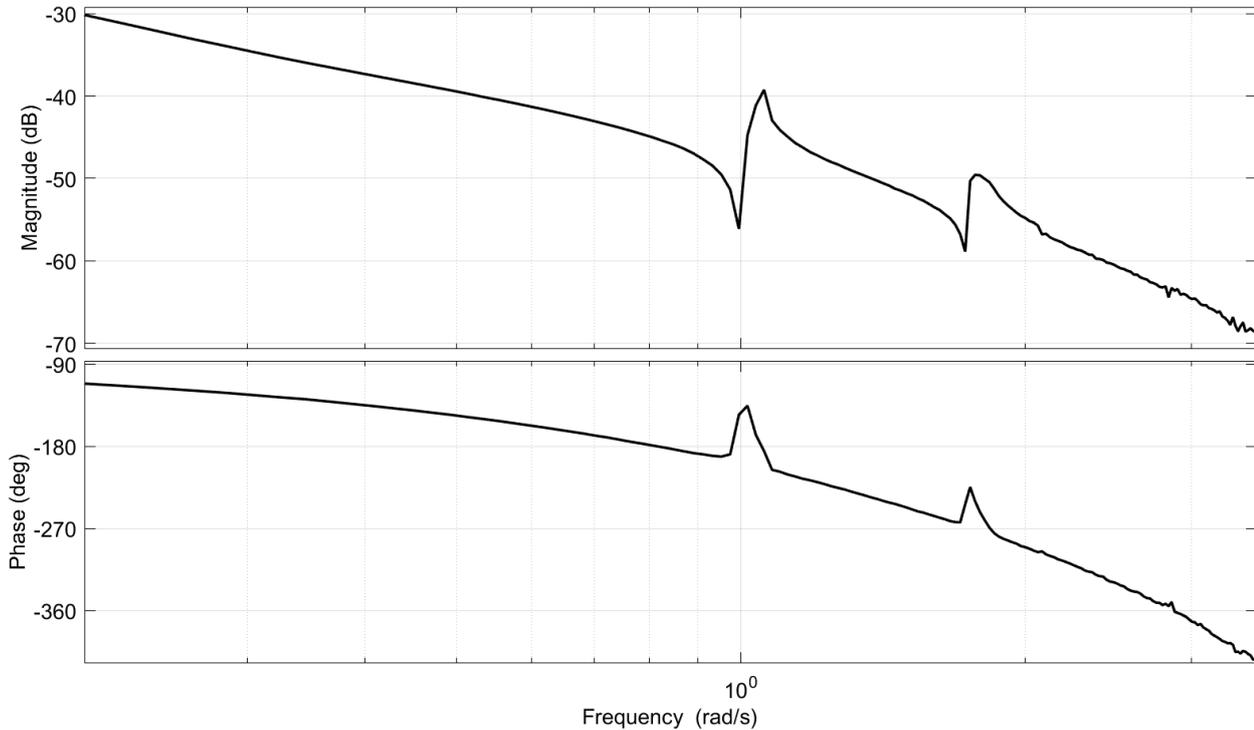


FIGURE 1.8 – Réponse fréquentielle de  $G$

Disposant d'un modèle du système à asservir, nous pouvons maintenant concevoir une loi de commande permettant de stabiliser celui-ci.

## 1.5 Choix d'une méthode de synthèse de loi de commande

Nous avons vu que la *modélisation* d'un système physique peut se faire de plusieurs manières : soit par la description de phénomènes physiques, soit par des mesures expérimentales. Dans les deux cas, un modèle sera donc toujours différent du système réel. Autrement dit, il existera toujours des *incertitudes*, des erreurs de modélisation, liées au fait que les phénomènes physiques ne peuvent être qu'imparfaitement traduits par des modèles mathématiques. Il faut donc accepter de travailler avec des modèles peu précis et imparfaits.

La *commande PID* [ÅH06, Bod45], apparue dans les années 50, effectue un compromis entre performances (stabilité, temps de réponse, bande passante) et contraintes de robustesse (*marges de gain* et de *phase*). Cependant, ces techniques se limitent à des systèmes *Simple Entrée Simple Sortie* (abrégé SISO pour *Simple Input Simple Output*) d'ordre très faible, et il est courant d'utiliser des techniques graphiques à l'aide des diagrammes de *Black-Nichols*,

*Nyquist* ou de *Bode* afin de régler manuellement les paramètres du correcteur. La commande PID a néanmoins mis en évidence le rôle fondamental de la boucle fermée par rapport à la boucle ouverte et a introduit des notions pratiques de robustesse (*marges de gain* et de *phase*, etc.).

La *commande optimale*, et plus particulièrement la *commande Linéaire Quadratique (LQ)*, est ensuite apparue dans les années 60 sous l'impulsion de Kalman [Kal60, Kal63, Kal64]. Ces méthodes de commande apparaissent en même temps que le concept de *représentations d'état* d'un système. L'objectif est de minimiser un critère de performance quadratique portant sur "l'énergie" de l'état d'un système et de sa commande. Le concepteur peut alors agir sur des matrices de pondération afin de régler la dynamique de la boucle fermée. La conception de ce type de commande se fait dans le domaine temporel, perdant ainsi toutes les notions introduites par l'étude des systèmes dans le domaine fréquentiel

La *commande  $H_\infty$*  est un concept apparu dans les années 80 [Zam81, ZF83, Vid85] qui permet d'approfondir ces contraintes de performance et de robustesse en boucle fermée. Celle-ci a pour objectif de concevoir des correcteurs satisfaisant certaines contraintes de stabilité et de robustesse en "modelant" des fonctions de transfert. Cette méthode permet l'utilisation des techniques de l'automatique fréquentielle classiques introduites par Bode, etc., ainsi que les méthodes de synthèse introduites par la commande optimale.

Il existe deux grandes approches dans la commande  $H_\infty$ . La première traite le problème dit *standard* où l'on cherche à minimiser la norme  $H_\infty$  de la fonction de transfert entre l'entrée et la sortie du système bouclé. L'avantage de cette formulation est la possibilité de prendre en compte des pondérations fréquentielles sur la boucle fermée, c'est-à-dire de modéliser directement certains transferts particuliers en boucle fermée. Ces modélisations peuvent se faire sur l'intégralité de la plage de fréquences, ce qui constitue un grand intérêt de la méthode. La résolution numérique d'un tel problème se fait classiquement à l'aide de deux méthodes. La première porte le nom d'*algorithme de Glover-Doyle* [DGKF89, GD88] et requiert la résolution d'*équations de Riccati algébriques* (ERA). Cet algorithme permet d'approximer la valeur optimale du critère  $H_\infty$  par dichotomie (on parle alors de  $\gamma$ -*itération*). L'autre méthode se base sur la résolution du *Lemme Borné Réel*. Celui-ci peut alors s'étudier à l'aide d'une *Inégalité Matricielle Linéaire* (LMI) [SW00, ACA+99, GA94].

Le deuxième approche traite le problème de commande dite par *Loop-Shaping* c'est-à-dire, en traduisant mot à mot, de commande par *modélage de la boucle ouverte*. Ce problème cherche à minimiser le maximum de la norme  $H_\infty$  d'une matrice formée par 4 fonctions de transferts particulières qui sont directement liés à la performance et la robustesse de l'asservissement. Le modelage de ces fonctions de transfert se fait en *boucle ouverte* en ajoutant des pondérations fréquentielles [MG92]. Le désavantage majeur qui en découle est que ces modelages ne peuvent se faire qu'à gain de boucle très faible ou très élevé. En revanche, l'avantage de cette méthode (et qui en fait sa force) réside dans la résolution du problème d'optimisation de la norme  $H_\infty$  : sa solution optimale est *explicite* [GM88]. Celle-ci s'exprime grâce à la résolution de deux *équations de Riccati algébriques*.

En pratique, la commande  $H_\infty$  a largement fait ses preuves dans le domaine de l'industrie [Mam92, Rah04, Cle15], notamment dans le domaine des viseurs [Cou97, Hir13]. En effet, le formalisme  $H_\infty$  permet d'introduire des critères de robustesse très pratiques pour gérer le vieillissement de produits ou les incertitudes de fabrication.

De plus, le réglage des paramètres du correcteur via des pondérations se fait de manière itérative. C'est souvent un processus long et fastidieux pour l'ingénieur automatique. Ceci devient un problème lorsque le nombre de projets qu'un ingénieur doit gérer devient important, et que les délais de conception sont de plus en plus courts afin de réduire les coûts. Notamment, au cours de la vie d'un même projet, l'architecture du système est susceptible de changer à de nombreuses reprises, et il est alors nécessaire de refaire des calculs de correcteurs. Aussi, lors d'essais en qualification, il apparaît souvent que les modèles prévus ne correspondent pas exactement avec le modèle réel, voire peuvent varier suivant les essais.

La *commande LPV (Linear Parameter Varying)* est une stratégie de synthèse allant dans ce sens. Elle peut se voir comme une extension de la commande  $H_\infty$  à des systèmes linéaires à paramètres variants. Elle apporte la robustesse de la méthode  $H_\infty$  et permet également au contrôleur de s'adapter à une variation connue (ou estimée) de paramètres [Sch01]. Elle repose néanmoins sur des variations paramétriques d'un modèle, c'est-à-dire sur le fait que la structure de celui-ci reste inchangée. De plus, l'aspect algorithmique de la synthèse reste complexe.

Il convient alors de chercher une alternative à ces techniques numériques afin de pouvoir recalculer des correcteurs de manière simple et efficace, tout en gardant un formalisme de commande robuste. Dans cette perspective, cette thèse a choisi de s'intéresser à la résolution du problème de *commande  $H_\infty$  par Loop-Shaping pour des systèmes à paramètres non fixés* (c'est-à-dire libres). Ce dernier problème a une solution explicitable via des solutions d'équations de Riccati algébriques. Le fait de pouvoir expliciter la solution d'un problème si complexe est une aubaine si l'on souhaite faire de la commande à paramètres inconnus. En effet, il n'est alors plus nécessaire d'utiliser des algorithmes d'optimisation pour minimiser la norme  $H_\infty$ . C'est en grande partie ce qui motiva le choix d'une telle méthode de commande. Celle-ci est à la fois simple à mettre en œuvre et prend en compte des critères de performance et de robustesse complexes.

## 1.6 État de l'art sur la résolution d'équations de Riccati algébriques et la commande $H_\infty$ paramétrique

Résoudre le problème de commande  $H_\infty$  par Loop-Shaping fait appel à la résolution d'*équations de Riccati algébriques* (ERA). Celle-ci se présente sous sa forme générale

$$X A + A^T X - X R X + Q = 0,$$

où  $X \in \mathbb{R}^{n \times n}$  est l'inconnue,  $A \in \mathbb{R}^{n \times n}$ ,  $Q = Q^T \in \mathbb{R}^{n \times n}$ ,  $Q > 0$  et  $R = R^T \in \mathbb{R}^{n \times n}$ ,  $R > 0$ . Cette équation n'est qu'une extension au cas matriciel de l'équation scalaire du second degré :

$$-r x^2 + 2 a x + q = 0.$$

Il n'y a généralement pas unicité de la solution  $X$  de l'ERA.

Afin de concevoir un correcteur  $H_\infty$  par Loop-Shaping pour un système à paramètres inconnus, il est nécessaire de résoudre une telle équation pour des matrices  $A$ ,  $R$  et  $Q$  qui dépendent de ces mêmes paramètres inconnus.

Les premiers travaux explorant ce problème furent menés par Forsman et Eriksson [FE93a, FE93b]. Ceux-ci considèrent l'ERA comme un système d'équations polynomiales  $\{e_{i,j} = 0 \mid (i,j) \in \llbracket 0, n \rrbracket^2\}$ , où  $e_{i,j}$  représente le coefficient de ligne  $i$  et la colonne  $j$  de l'ERA de taille  $n \times n$ . A ce système, ils appliquèrent des techniques de calcul formel (*bases de Gröbner* [CLO15, Chap. 2]) afin de réduire ce système à un système plus simple. Leurs travaux restèrent très empiriques, Forsman et Eriksson déclarant que plusieurs questions restaient ouvertes, notamment “Est-il possible de réduire le système à une seule équation polynomiale univariée (c'est-à-dire pour une équation de n'importe quelle taille, et avec n'importe quel nombre de paramètres) ?” ou encore “Est-il toujours possible de réduire le nombre d'équations non-linéaires en mettant le système sous forme canonique ?”.

Ces travaux furent poursuivis dans [KYAH09], où il fut notamment montré que l'ERA peut se ramener, dans le cas générique, à un système de  $n$  équations polynomiales  $\mathcal{B}$ ,  $n$  désignant la taille des matrices considérées. Ce système d'équations polynomiales est dérivé d'une *factorisation spectrale* [AHKY09, JK85] de la *matrice Hamiltonienne* [ACA<sup>+</sup>99] associée à l'ERA. Plus particulièrement, il fut montré que ce système de  $n$  équations possède des propriétés bien particulières (*base de Gröbner réduite* [CLO15, Chap. 2]) permettant d'appliquer des techniques d'éliminations de variables. Notamment, [KYAH09] prouve que ce système est sous forme *shape* dans le cas générique, c'est-à-dire, grossièrement, que les solutions de celui-ci sont paramétrables par une seule variable  $t \in \mathbb{C}$  vérifiant une équation polynomiale  $p(t) = 0$ . Réutilisant leurs résultats sur la factorisation spectrale [AHY05], [KYAH09] montre que la solution définie positive de l'ERA est définie par la racine réelle maximale du polynôme univarié  $p$ . Ce dernier résultat permet de s'affranchir du *critère de Sylvester* [Gil91] pour identifier une matrice définie positive, critère portant sur l'étude du signe des  $n$  mineurs principaux dominants de la matrice. Ce résultat permet donc de transformer un critère semi-algébrique donné par les conditions de positivité des mineurs principaux dominants, en un “critère algébrique” donné par la solution réelle maximale d'une équation polynomiale.

Ce même groupe de recherche proposa ensuite d'étendre leurs travaux sur l'ERA à la commande  $H_\infty$ . Dans [KHO07], ce groupe présente des expressions explicites de critères  $H_\infty$  pour des systèmes particuliers d'ordre faible, ainsi que des correcteurs explicites pour des systèmes de type  $1/s^n$ . Ce dernier travail permet de caractériser des systèmes facilement commandables. Aussi, [KH12] propose une méthode pour concevoir un correcteur  $H_\infty$  atteignant un critère  $H_\infty$  minimal. Pour cela, ils utilisent des pondérations fréquentielles comme moyen de réglage et un algorithme de descente afin de converger vers un minimum local du critère dans l'espace des paramètres. Par ailleurs, [Muh07] propose l'utilisation de méthodes graphiques pour la conception de correcteurs robustes, notamment  $H_\infty$ .

De manière générale, à notre connaissance, peu de travaux mêlant calcul formel et automatique existent dans la littérature. En effet, il est d'usage de résoudre ces problèmes d'automatique complexes en utilisant des algorithmes d'optimisation numériques. L'avantage est alors de pouvoir travailler avec des *modélisations fines* prenant en compte de nombreux phénomènes pratiques. Cette thèse propose une alternative en promouvant l'aspect formel dans les applications de l'automatique et du contrôle. L'objectif recherché est de permettre d'appréhender les problèmes de l'automatique linéaire d'une *manière nouvelle*, pour des systèmes définis par des *modélisations simples*. Cette alternative se veut *complémentaire* des méthodes numériques usuelles, donnant en cela de nouveaux outils à l'automaticien.

## 1.7 Plan de thèse

Dans la continuité de ces travaux, dans ce mémoire, nous proposons d'étudier la commande  $H_\infty$  à l'aide des techniques de calcul formel. Dans la rédaction de ce mémoire, nous avons essayé d'apporter le plus grand soin à ce qu'il soit accessible tant à la communauté automatique qu'à celle des algébristes/calcul formel. Pour cela, nous illustrerons les nombreux concepts par des exemples en utilisant un vocabulaire simple et sans rentrer dans des détails trop techniques. Cependant, les résultats théoriques de cette thèse seront tous prouvés rigoureusement, et nous nous référerons parfois à des détails techniques mis en annexe. La démarche que nous proposons est la suivante :

- Le Chapitre 2 reprend les fondements de la commande  $H_\infty$  par Loop-Shaping et montre les différentes facettes de ce problème. Ainsi, le lecteur pourra se familiariser avec le formalisme de l'automatique et comprendre comment est classiquement abordé ce problème. Ce problème de commande s'étudie à l'aide de deux ERA. Cependant, dans le cas des applications aux viseurs, nous considérons souvent des systèmes SISO. Ainsi, dans le cas de systèmes SISO sans transfert direct, nous montrerons que ce problème se simplifie et ne nécessite que la résolution d'une seule ERA notée  $\mathcal{R} = 0$ . Nous rappellerons aussi les liens étroits entre commande  $H_\infty$  et commande LQG. Nous finirons par un exemple d'application pratique à un modèle de viseur.
- Au Chapitre 3, nous montrons comment réduire l'ERA à un système d'équations polynomiales par deux méthodes différentes. La première inspectant directement le système d'équations polynomiales découlant de l'ERA dans le cas où le système est considéré sous forme canonique. Nous remarquons alors un schéma de substitutions simplifiant grandement les calculs, et ce à n'importe quel ordre. La deuxième, reprenant l'approche de [KYAH09], se base sur la recherche d'une factorisation spectrale de la matrice Hamiltonienne associée à l'ERA, ainsi qu'une paramétrisation des vecteurs propres de cette matrice. Nous montrons alors l'équivalence des systèmes obtenus par les deux approches. En Section 3.4, nous exploitons ces propriétés à l'aide d'outils de la théorie de l'élimination afin d'explicitier une paramétrisation de ce système. Nous proposons alors une paramétrisation rationnelle, la *Représentation Rationnelle Univariée* [Rou99, Rou07], présentant de nets avantages par rapport à d'autres paramétrisations (telle qu'une *base de Gröbner sous forme shape* [CLO15, App. D.2]). En reprenant les travaux de [KYAH09], nous montrons ensuite comment isoler l'unique matrice définie positive solution de l'ERA. Nous explicitons ensuite des formes closes pour la solution de  $\mathcal{R} = 0$  pour  $1 \leq n \leq 4$ . Finalement, nous détaillons quelques cas particuliers remarquables.
- Au Chapitre 4, nous étendons l'étude paramétrique de l'ERA à la commande  $H_\infty$  par Loop-Shaping. Cette extension passe par le calcul de valeurs propres d'une matrice construite à partir de la solution de l'ERA. Ce calcul de valeurs propres est réalisé à l'aide d'un calcul de polynôme caractéristique. Lorsque les polynômes sont d'ordre suffisamment faible, nous verrons que nous pouvons expliciter les racines par radicaux et identifier des racines maximales à l'aide du concept de *variété discriminante*. En revanche, quand ce n'est pas le cas, une étape numérique est alors nécessaire. Nous détaillons les deux approches dans ce chapitre : l'approche symbolique en petite dimension et l'approche symbolique/numérique dans les autres cas. Nous présentons aussi une méthode pour régler une pondération Loop-Shaping scalaire de manière optimale

sur deux exemples (un système deux masses ressorts non amorti et amorti). Celle-ci se base sur le calcul d'un polynôme encodant toutes les valeurs propres de toutes les matrices solutions des équations de Riccati intervenant dans le problème  $H_\infty$ .

- Au Chapitre 5, nous étendrons nos résultats obtenus en commande  $H_\infty$  paramétrique aux systèmes à retard. Pour cela, nous utiliserons une décomposition additive du système avec retard. Celle-ci nous permettra de concevoir des correcteurs explicites stabilisant le système avec retard à partir de correcteurs stabilisant le système sans retard. Nous terminerons en proposant une méthode de réglage de pondération optimale pour ces systèmes avec retard, intéressante en pratique pour les systèmes considérés chez Safran Electronics & Defense.
- Enfin, la conclusion permet de résumer les résultats obtenus dans cette thèse et de formuler de nombreuses perspectives venant de l'utilisation des méthodes du calcul formel en automatique et, en particulier, en commande  $H_\infty$ .

## 1.8 Liste de publications

### Conférences internationales :

- Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat. A symbolic-numeric method for the parametric  $H_\infty$  loop-shaping design problem, 22<sup>nd</sup> *International Symposium on Mathematical Theory of Networks and Systems*, 2016.

*Dans ce papier, nous présentons une méthode symbolique-numérique pour résoudre le problème de commande  $H_\infty$  par Loop-Shaping pour des systèmes LTI, SISO, d'ordre faible avec paramètres. La présence de paramètres inconnus ne permet plus l'utilisation directe des méthodes numériques classiques pour la résolution du problème de commande  $H_\infty$  et nécessite donc le développement de nouveaux algorithmes. En utilisant les concepts de base de Gröbner et de Représentation Univariée Rationnelle de variétés algébriques zéro-dimensionnelles, nous donnons une paramétrisation de toutes les solutions des deux équations de Riccati algébriques associées au problème de contrôle  $H_\infty$ . Ensuite, en utilisant un résultat de [AHKY09] sur la factorisation spectrale, nous obtenons un algorithme symbolique-numérique certifié pour calculer les solutions définies positives des équations de Riccati algébriques. L'approche est enfin illustrée par un exemple standard en automatique.*

- Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat. Explicit  $H_\infty$  controllers for 1<sup>st</sup> to 3<sup>rd</sup> order single-input single-output systems with parameters, 20<sup>th</sup> *International Federation on Automatic Control World Congress*, 2017.

*Dans ce papier, à l'aide de méthodes de calcul formel pour l'étude des systèmes polynomiaux, nous calculons sous forme close des correcteurs  $H_\infty$  pour des systèmes LTI, SISO, d'ordre inférieur ou égal à 3 avec paramètres inconnus.*

- Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat, Fabrice Rouillier. Explicit  $H_\infty$  controllers for 4<sup>th</sup> order single-input single-output systems with parameters and their application to the two mass-spring system with damping, 20<sup>th</sup> *International Federation on Automatic Control World Congress*, 2017.

*Dans ce papier, à l'aide de méthodes de calcul formel pour l'étude des systèmes polynomiaux, nous calculons sous forme close des correcteurs  $H_\infty$  pour des systèmes LTI, SISO, d'ordre 4 avec paramètres inconnus. Nous présentons aussi une méthode de réglage optimal d'une pondération Loop-Shaping scalaire de type gain pour un exemple standard de l'automatique : le système masse-ressort.*

- Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat. Parametric sub-optimal  $H_\infty$  controllers for an optro-mechanical system modeled by a time-delay 4<sup>th</sup> order system, 8<sup>th</sup> *International Symposium on Optronics in Defence and Security*, 2018.

*Dans ce papier, nous étendons les correcteurs  $H_\infty$  paramétriques sous forme close développés dans les articles précédents à une classe de systèmes à retard dont la partie sans retard est d'ordre inférieur ou égal à 4. Cette extension est réalisée en utilisant une décomposition additive du système en une partie sans retard et une partie avec un retard distribué. Nous explicitons ensuite une démarche de réglage de pondération Loop-Shaping scalaire pour cette classe de systèmes à retard et nous illustrons nos résultats sur un modèle de viseur gyro-stabilisé.*

- Axel Barrau, Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat. Using symbolic computation to solve algebraic Riccati equations arising in invariant filtering, 17<sup>th</sup> *European Control Conference*, 2018.

*Dans le cadre de la navigation inertielle, ce papier propose d'utiliser la méthode symbolique pour la commande  $H_\infty$  paramétrique, développée dans les papiers précédents, pour caractériser la solution définie positive d'une équation de Riccati algébrique venant d'un problème de filtrage invariant.*

#### Revue scientifique :

- Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat. A parametric robust control approach for low order systems, *en cours de soumission*.

*Ce papier fait une synthèse de l'ensemble des résultats pour la commande  $H_\infty$  paramétrique obtenus dans les papiers de conférence précédents et dans ce mémoire de thèse.*

# Chapitre 2

## Fondements de la commande $H_\infty$ par Loop-Shaping

L'introduction de cette thèse a motivé le choix de la commande  $H_\infty$  par Loop-Shaping comme méthode de commande pour l'asservissement des viseurs développés par *Safran Electronics & Defense*. Cette méthode est basée sur le modelage de la boucle ouverte. Ce modelage se fait dans le domaine fréquentiel à l'aide de pondérations.

Afin de comprendre ces concepts de modelage fréquentiels, la première section de ce chapitre détaille la signification physique des 4 fonctions de sensibilité, centrales dans ce problème de commande. Dans un second temps, nous expliquons comment modeler ces fonctions de sensibilité par des considérations sur la boucle ouverte. La Section 2.3 détaille ensuite l'intérêt de la commande  $H_\infty$  et les résultats fondamentaux d'un problème  $H_\infty$  particulier, originellement présentés dans [GM88]. La Section 2.5 effectue le lien entre ce problème  $H_\infty$  et le concept de Loop-Shaping en reprenant la méthode de synthèse proposée dans [MG92]. La Section 2.4 montre un lien entre la commande  $H_\infty$  et la commande *Linéaire Quadratique Gaussienne*. Enfin, cette théorie est appliquée à une modélisation de viseur.

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>30</b>
<b>2.2</b>	<b>Caractérisation fréquentielle des objectifs de performance et de robustesse</b>	<b>32</b>
2.2.1	Fonction de sensibilité et contraintes de performance	32
2.2.2	Fonction de sensibilité complémentaire et contraintes de robustesse	32
2.2.3	Fonction de régulation et contraintes de performance en régulation	34
2.2.4	Fonction de commande et contraintes de saturation	34
<b>2.3</b>	<b>Stabilisation de systèmes aux facteurs copremiers incertains via critère <math>H_\infty</math></b>	<b>35</b>
2.3.1	Factorisations copremières d'une matrice de transfert	35
2.3.2	Un problème de commande robuste	36
2.3.3	Conception d'un correcteur	38
2.3.4	Cas particulier des systèmes SISO	40
<b>2.4</b>	<b>Lien avec la commande Linéaire Quadratique Gaussienne</b>	<b>43</b>
<b>2.5</b>	<b>Commande <math>H_\infty</math> par Loop-Shaping</b>	<b>44</b>
2.5.1	Modelage de la boucle ouverte	44
2.5.2	Une procédure de synthèse de correcteur $H_\infty$ par Loop-Shaping	47
<b>2.6</b>	<b>Exemple d'application</b>	<b>48</b>
<b>2.7</b>	<b>Conclusion</b>	<b>52</b>

---

## 2.1 Introduction

Dans ce chapitre, nous considérons un système linéaire  $G$ , de dimension finie et invariant dans le temps. Nous désignons par  $(A, B, C, D)$  une représentation d'état minimale de  $G$ , et par  $G = C(sI - A)^{-1}B + D$  sa matrice de transfert. La classe de systèmes considérée en pratique étant strictement propre, nous considérerons que  $G$  est de transfert direct nul, c'est-à-dire  $D = 0$ . Nous noterons alors simplement  $(A, B, C)$  une représentation d'état de  $G$ . Ceci nous permet aussi d'alléger les résultats théoriques classiques présentés dans ce chapitre, faisant intervenir beaucoup plus de termes dans le cas  $D \neq 0$ . Par ailleurs, nous supposons la paire  $(A, B)$  commandable et  $(C, A)$  observable (nous parlons aussi de *représentation minimale* de  $G$ ). Nous notons aussi  $I_n$  la matrice identité de taille  $n \times n$ , ou simplement  $I$  si la dimension de celle-ci est évidente. De même, nous noterons par  $\mathbb{O}_{m \times p}$  la matrice de taille  $m \times p$  composée uniquement de zéros.

Dans un souci de lisibilité, nous nous limiterons à l'étude de systèmes Simple Entrée Simple Sortie (abrégé par SISO pour *Simple Input Simple Output*) pour expliquer le concept de Loop-Shaping. Remarquons cependant que ce concept s'étend aux systèmes Multi-Entrées Multi-Sorties (abrégé par MIMO pour *Multiple Input Multiple Output*) via des considérations de *valeurs singulières*. Le lecteur peut se référer, par exemple, à [DF00] pour une description rapide et, par exemple, à [GM89] pour plus de détails.

Nous notons par  $H_\infty$  l'*algèbre de Hardy* formée par les fonctions holomorphes dans  $\mathbb{C}_+ := \{s \in \mathbb{C} \mid \text{Re}(s) > 0\}$  bornées pour la norme  $\|G\|_\infty = \sup_{s \in \mathbb{C}_+} |G(s)| = \text{ess sup}_{\omega \in \mathbb{R}} |G(i\omega)|$  [Zam81, ZDG96]. Nous désignons par  $RH_\infty$  le sous-espace de  $H_\infty$  constitué des fonctions rationnelles bornées et holomorphes dans  $\mathbb{C}_+$ , c'est-à-dire  $RH_\infty := \mathbb{R}(s) \cap H_\infty$ . En d'autres termes,  $RH_\infty$  est l'algèbre des fractions rationnelles propres et stables.

Nous considérons un système  $G$  de commande  $u$  et de mesure  $y$ , introduit en Section 1.3, bouclé par un correcteur  $K$  donné par le schéma de la Figure 2.1. Nous voulons faire tendre la sortie  $y$  vers une référence  $r$ , bien que le système soit soumis à des perturbations  $d$  en entrée de système et  $\delta$  en sortie.

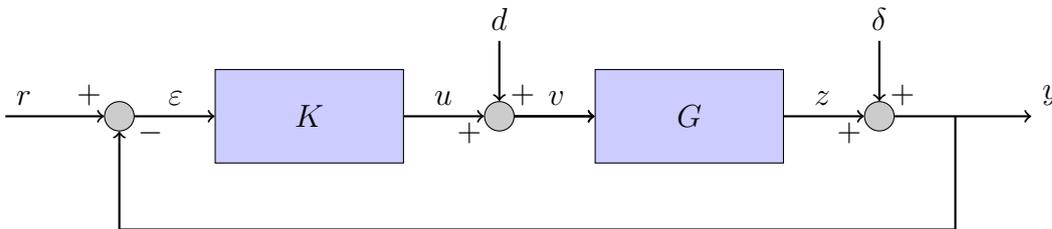


FIGURE 2.1 – Représentation d'un système bouclé

Nous considérons des fonctions de transfert  $G$  et  $K$  respectivement du système et du correcteur. Définissons la boucle ouverte et sa bande passante, notions que l'on réutilisera par la suite.

**Définition 2.1.** La *boucle ouverte* est la fonction de transfert définie par :

$$L := GK.$$

Sa *bande passante* à 0 dB, notée  $f_0$  en Hertz ou  $s_0 := 2i\pi f_0$  en radian, est la plus petite pulsation positive telle que :

$$L(s_0) = G(s_0)K(s_0) = 1.$$

Nous définissons aussi la fonction suivante qui interviendra de manière récurrente dans ce chapitre.

**Définition 2.2.** La **fonction de sensibilité**, notée  $S$ , est définie par :

$$S := (1 + GK)^{-1}.$$

En s'appuyant sur le schéma de la Figure 2.1, nous déduisons :

$$y = \delta + z = \delta + Gv = \delta + Gd + Gu = \delta + Gd + GK\varepsilon = \delta + Gd + GKr - GK y.$$

Sous réserve que  $1 + GK$  est inversible (la boucle fermée est alors dite *bien posée*), nous obtenons :

$$y = S\delta + GSd + GKSr.$$

De plus, en utilisant  $u = K\varepsilon = Kr - Ky$ , nous avons :

$$u = -KS\delta - GKSd + K Sr.$$

Nous avons alors :

$$\begin{pmatrix} y \\ -u \end{pmatrix} = \begin{pmatrix} S & GS & GKS \\ KS & GKS & -KS \end{pmatrix} \begin{pmatrix} \delta \\ d \\ r \end{pmatrix}.$$

Notons que cette écriture est indépendante de la représentation des systèmes  $G$  et  $K$ . Dans le cadre de la stabilisation de LdV, nous avons une consigne/référence nulle ( $r = 0$ ). Nous obtenons alors le schéma de la Figure 2.2. En introduisant un signe  $-$  dans la boucle interne et en posant  $\mu = -u$ , le schéma de la Figure 2.2 se simplifie légèrement par rapport à la situation précédente. C'est ce type de schéma qui est considéré dans la littérature sur la commande  $H_\infty$ . Remarquons alors que  $\delta$  a une double signification d'un point de vue transfert : il peut représenter une perturbation ou une référence car le transfert entre  $\delta$  et  $u$  (resp.  $y$ ) est similaire au transfert entre  $r$  et  $u$  (resp.  $y$ ).

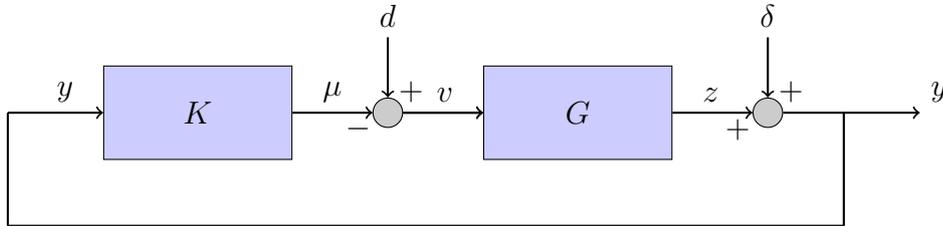


FIGURE 2.2 – Boucle de stabilisation

Nous définissons aussi la notion de stabilité suivante.

**Définition 2.3.** Le système bouclé de la Figure 2.2 est dit *stable de manière interne* si les matrices de transfert  $S$ ,  $KS$ ,  $GS$  sont propres et stables, c'est-à-dire appartiennent à  $RH_\infty$ . Dans ce cas,  $G$  est dit *stabilisable de manière interne* et  $K$  est appelé *correcteur stabilisant*.

Dans la suite, nous adoptons la notation suivante :

$$\Pi(K, G) := \begin{pmatrix} S & SG \\ KS & KSG \end{pmatrix} = \begin{pmatrix} I \\ K \end{pmatrix} S \begin{pmatrix} I & G \end{pmatrix}. \quad (2.1)$$

Remarquons que  $\Pi(K, G)^2 = \Pi(K, G)$ . Nous avons alors le lemme suivant.

**Lemme 2.1.**  $K$  stabilise  $G$  de manière de interne si et seulement si  $\Pi(K, G)$  est un projecteur stable, c'est-à-dire  $\Pi(K, G) \in RH_\infty$ .

Par ailleurs, nous venons d'introduire quatre fonctions  $S$ ,  $GS$ ,  $KS$  et  $KGS$ . Celles-ci permettent de caractériser la boucle de stabilisation (2.2) en terme de performance et de robustesse. La prochaine section détaille la signification pratique de chacune de ces fonctions.

## 2.2 Caractérisation fréquentielle des objectifs de performance et de robustesse

### 2.2.1 Fonction de sensibilité et contraintes de performance

Nous avons défini la fonction de sensibilité comme étant  $S := (I + GK)^{-1}$ . Elle peut être aussi perçue comme le transfert de l'atténuation de la perturbation en sortie sur la mesure :

$$\frac{y}{\delta} = S = (I + L)^{-1}.$$

Ainsi,  $S$  est un indice de performance de la boucle d'asservissement en *rejet de la perturbation de sortie*. On cherchera à annuler l'erreur en basses fréquences (pour  $f < f_0$ ), c'est-à-dire à avoir  $|S| \ll 1$  en basses fréquences (voir la Section 2.5.1 pour la justification). La forme de  $S$  recherchée est donc de type *passé haut* avec en plus des *réjecteurs* sur les fréquences à atténuer.

### 2.2.2 Fonction de sensibilité complémentaire et contraintes de robustesse

Supposons que le modèle (nominal), noté  $G_0$ , s'écarte du système réel  $G$  par des *erreurs multiplicatives*  $\Delta$  :

$$G_0 = (1 + \Delta) G.$$

Dans le cas d'un système SISO, pour garantir la stabilité pour une telle incertitude de modélisation, il suffit que l'écart entre le lieu de la fonction de transfert en boucle ouverte nominale et le point critique  $(-1, 0)$  soit supérieur à l'écart maximal entre la boucle ouverte réelle et la boucle ouverte nominale (Figure 2.3). En effet, d'après le *critère de Nyquist* [Vin00, p.17], dépasser le point critique entraîne une perte de stabilité BIBO (c'est-à-dire pour une entrée  $u$  bornée, la sortie  $z$  n'est pas bornée).

Traduisons ce phénomène sous forme d'inégalités. L'écart de module entre les deux boucles ouvertes est donné par  $|K(G_0 - G)| = |K\Delta G|$ . Il suffit que cet écart soit inférieur à la distance entre la boucle ouverte nominale et le point critique  $|1 + KG|$ . Ceci se traduit par l'inégalité suivante :

$$|K\Delta G| < |1 + KG| \quad \Leftrightarrow \quad |LS| = \left| \frac{KG}{1 + KG} \right| < \frac{1}{|\Delta|}. \quad (2.2)$$

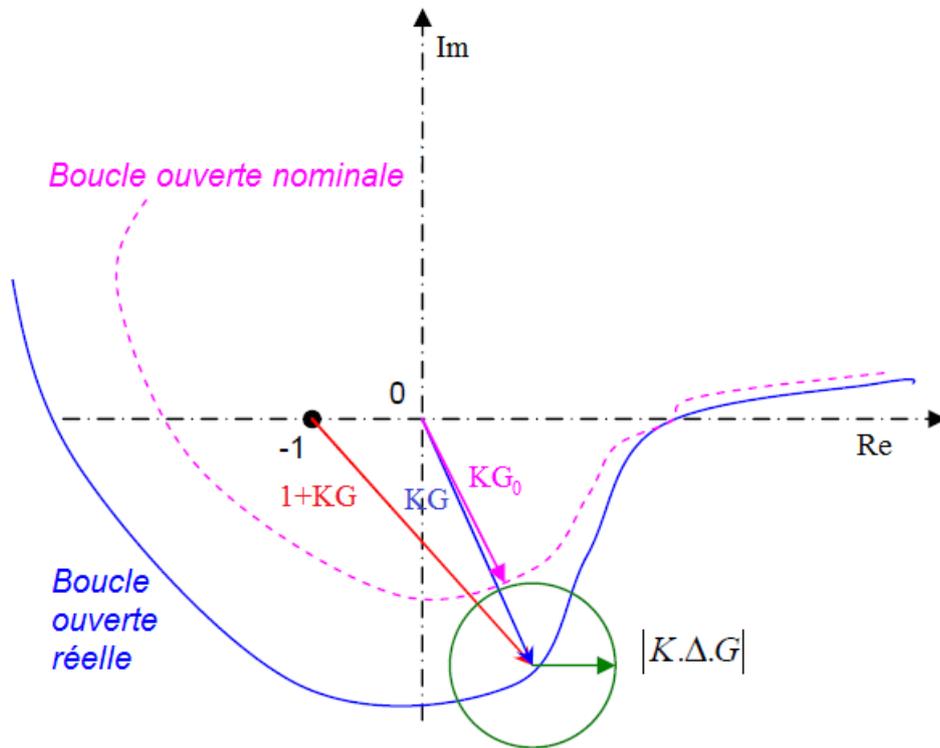


FIGURE 2.3 – Incertitudes de modélisation

Cette inégalité fait directement intervenir  $LS$ , que l'on nomme *fonction de sensibilité complémentaire*. Cette fonction est donc un indice de la *robustesse en stabilité pour les erreurs multiplicatives* de la boucle. Elle prend en compte les variations du modèle qui sont souvent plus importantes en haute fréquences.

Cette fonction représente aussi le transfert entre la commande et la perturbation en entrée de modèle :

$$\frac{\mu}{d} = LS.$$

En basse fréquence, celle-ci tendra vers 1, ce qui signifie que la perturbation sera compensée par la commande. En haute fréquence, les erreurs de modèle rendront cette tâche difficile. Nous chercherons alors à incorporer des comportements de type filtres passes-bas dans cette fonction (voir la Section 2.5.1 pour la justification).

**Remarque 2.1.** Il est important de noter que  $S + LS = 1$ . Ceci traduit le *compromis* inévitable entre la performance (critère sur  $S$ ) et la robustesse (critère sur  $LS$ ), inhérent à toute méthode de contrôle [ZDG96].

**Remarque 2.2.** En pratique, les modes de résonance mécaniques sont susceptibles de changer de quelques pour-cents d'un viseur à l'autre. Pour un même viseur, le(s) mode(s) de résonance dépend(ent) aussi du niveau d'excitation et de température. Les modes de résonance sont donc une limitation à la robustesse de l'asservissement et donc à sa performance.

**Remarque 2.3.** Noter que le raisonnement ci-dessus s'étend aux systèmes MIMO à l'aide du *théorème du Petit Gain* [ZDG96, p. 211][Zam81].

### 2.2.3 Fonction de régulation et contraintes de performance en régulation

Considérons le transfert entre la perturbation en entrée de modèle  $d$  et la mesure  $y$  :

$$\frac{y}{d} = G S.$$

Ce transfert est appelé *fonction de régulation*. Dans le cadre des viseurs, un objectif de l'asservissement est d'avoir un mouvement résiduel le plus faible possible malgré les perturbations  $d$  (des couples de frottement). Ceci est en général envisageable à basse fréquence (voir la Section 2.5.1 pour la justification).

**Remarque 2.4.** Supposons que la boucle donnée à la Figure 2.2 décrive une boucle de vitesse d'un viseur. Nous voulons que l'erreur de position  $\vartheta(s) := \frac{y(s)}{s}$  soit nulle, bien que le viseur soit soumis à des couples perturbateurs  $d$ . Ces couples sont de type échelon d'amplitude  $C_0$ , soit  $d(s) = \frac{C_0}{s}$ . De plus, supposons que le correcteur comporte  $i$  intégrateurs :

$$K(s) = \frac{K_0(s)}{s^i}, \quad K_0(0) \neq 0.$$

Sous forme d'équations, cela se traduit par :

$$\begin{aligned} \lim_{t \rightarrow +\infty} \vartheta(t) &= \lim_{s \rightarrow 0} s \vartheta(s) && \text{(th. de la valeur finale)} \\ &= \lim_{s \rightarrow 0} s \frac{y(s)}{s} && \text{(déf. de l'erreur en vitesse)} \\ &= \lim_{s \rightarrow 0} G(s) S(s) d(s) && \text{(déf. fonction de régulation)} \\ &= \lim_{s \rightarrow 0} G(s) \frac{1}{1 + G(s) \frac{K_0(s)}{s^i}} \frac{C_0}{s} && \text{(perturbation type échelon)} \\ &= \lim_{s \rightarrow 0} \frac{s^i}{K_0(s)} \frac{C_0}{s} && \left( \text{approx. basse fréquence } \left| G(s) \frac{K_0(s)}{s^i} \right| \gg 1 \right) \\ &= \frac{C_0}{K_0(0)} \lim_{s \rightarrow 0} s^{i-1}. \end{aligned}$$

Ainsi, pour avoir une erreur de position angulaire nulle, c'est-à-dire  $\lim_{t \rightarrow +\infty} \vartheta(t) = 0$ , quand le viseur est soumis à des couples de type échelon, il faut que le correcteur comporte au moins deux intégrateurs ( $i \geq 2$ ).

### 2.2.4 Fonction de commande et contraintes de saturation

Dans le cadre des viseurs, du bruit dans le capteur ou une forte déformation peut faire saturer les amplificateurs de puissance. Une fois en saturation, la stabilisation est fortement dégradée car on quitte alors le régime linéaire. Dans ce cas, la stabilité de la boucle doit être étudiée d'un point de vue des systèmes non linéaires. Il est donc utile de connaître la *fonction de commande* définie par :

$$\frac{\mu}{\delta} = K S.$$

Une grande bande passante donne un correcteur avec un fort gain et peut amplifier les bruits au niveau des amplificateurs de puissance. Il est donc nécessaire de filtrer les hautes fréquences en incorporant des actions filtrantes dans  $K$  (voir la Section 2.5.1 pour la justification).

Nous venons de décrire l'importance de chacune de ces quatre fonctions dans la caractérisation d'une boucle de stabilisation. La prochaine section présente un problème de commande robuste dans lequel ces quatre fonctions sont centrales.

## 2.3 Stabilisation de systèmes aux facteurs copremiers incertains via critère $H_\infty$

Durant les années 80, la nécessité de prendre en compte des incertitudes de modèle dans la conception d'un correcteur est devenue incontournable. Les premiers travaux portèrent sur des incertitudes de type additive ou multiplicative [DS81, CD82]. Cependant, de telles incertitudes ne pouvaient prendre en compte toutes les variations possibles de modèles (comme par exemple le passage de modes stables à instables). C'est ce qui a amené Vidyasagar à introduire le concept d'incertitude sur les facteurs copremiers [VSF82]. Cette modélisation particulière de l'incertitude fut reprise par [GM89] dans le cadre de la stabilisation robuste d'une famille de systèmes.

Cette section reprend les résultats fondamentaux de ces travaux dans le cas des systèmes MIMO, puis montre des simplifications dans le cas SISO. Ces résultats portent sur des *systèmes linéaires*, notés  $G$ , de *dimension finie, invariants dans le temps*. Pour un rappel sur les systèmes linéaires et les définitions liées, le lecteur est invité à se reporter à l'Annexe A.

### 2.3.1 Factorisations copremières d'une matrice de transfert

Les résultats de cette section portent sur la prise en compte d'incertitudes sur les facteurs copremiers. Tout d'abord, définissons donc une *factorisation copremière* d'une matrice de transfert.

**Définition 2.4.** Soit  $G \in \mathbb{R}(s)^{q \times p}$ . Une *factorisation copremière à gauche* de  $G$  est un couple  $(\widetilde{M}, \widetilde{N}) \in RH_\infty^{q \times (q+p)}$ , avec  $\det(\widetilde{M}) \neq 0$ , et tel que :

- $G = \widetilde{M}^{-1} \widetilde{N}$ ,
- Il existe un couple  $(U, V) \in RH_\infty^{(p+q) \times q}$  tel que :  $\widetilde{M} V + \widetilde{N} U = I$ .

**Remarque 2.5.** On définit de manière duale une *factorisation copremière à droite*.

Il existe une infinité de telles représentations. Afin d'en distinguer une unique, nous pouvons considérer la propriété de normalisation suivante.

**Définition 2.5.** Une factorisation copremière à gauche  $(\widetilde{M}, \widetilde{N}) \in \mathbb{R}(s)^{q \times p}$  est dite *normalisée* si

$$\widetilde{M} \widetilde{M}^* + \widetilde{N} \widetilde{N}^* = I,$$

où  $\widetilde{M}^*(s) = \widetilde{M}^T(-s)$  (resp.  $\widetilde{N}^*(s) = \widetilde{N}^T(-s)$ ).

Une telle factorisation copremière normalisée de  $G$  peut se calculer à l'aide de la résolution d'une *équation de Riccati algébrique* [GM88], d'où son intérêt en pratique.

**Lemme 2.2.** [GM88] Considérons l'équation de Riccati suivante :

$$\widetilde{\mathcal{R}} := AY + Y A^T - Y C^T C Y + B B^T = 0.$$

Une factorisation copremière normalisée à gauche de  $G$  est alors définie par

$$\begin{cases} \tilde{N} := C (sI - A + Y C^T C)^{-1} B, \\ \tilde{M} := I + (sI - A + Y C^T C)^{-1} Y C^T, \end{cases}$$

où  $Y$  est l'unique solution définie positive de  $\tilde{\mathcal{R}} = 0$ .

### 2.3.2 Un problème de commande robuste

Soit  $(\tilde{M}, \tilde{N}) \in RH_\infty^{q \times (q+p)}$  une factorisation copremière à gauche normalisée de  $G$ . Nous pouvons alors définir une famille de systèmes à partir d'incertitudes  $(\Delta_M, \Delta_N)$  sur ses facteurs copremiers :

$$\mathcal{G}_\varepsilon = \left\{ G_\Delta := (\tilde{M} + \Delta_M)^{-1} (\tilde{N} + \Delta_N) \mid \|(\Delta_M \ \Delta_N)\|_\infty < \varepsilon \right\}. \quad (2.3)$$

Le problème de commande robuste est alors de trouver un correcteur sous forme de feedback stabilisant la famille de systèmes  $\mathcal{G}_\varepsilon$  (Figure 2.4). La propriété du système bouclé sera appelée *stabilité robuste* [GM89].

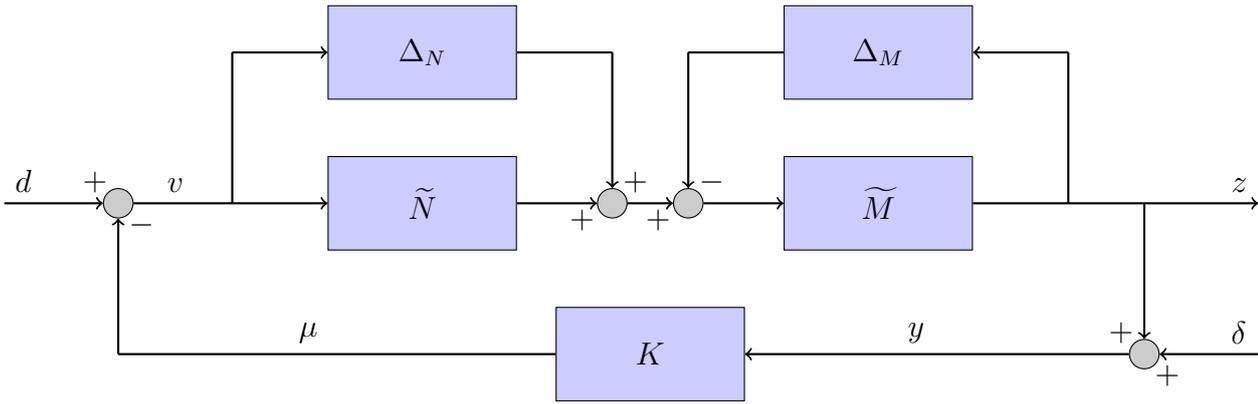


FIGURE 2.4 – Stabilisation d'une famille de modèles  $\mathcal{G}_\varepsilon$

**Définition 2.6.** Le système bouclé est dit *robustement stable* s'il est stable de manière interne pour tout  $G_\Delta$  dans  $\mathcal{G}_\varepsilon$ .

Nous avons alors le théorème suivant.

**Théorème 2.1.** [GM89] Le correcteur  $K$  stabilise de manière robuste le système  $G$  si et seulement si  $K$  stabilise  $G$  de manière interne et :

$$\left\| \begin{pmatrix} I \\ K \end{pmatrix} (I - GK)^{-1} \tilde{M}^{-1} \right\|_\infty \leq \gamma := \varepsilon^{-1}. \quad (2.4)$$

De manière équivalente, il existe un correcteur  $K$  tel que le système bouclé de la Figure 2.4 soit *robustement stable* si  $K$  stabilise  $G$  de manière interne et

$$\min_K \left\| \begin{pmatrix} I \\ K \end{pmatrix} (I - GK)^{-1} \tilde{M}^{-1} \right\|_\infty \leq \gamma := \varepsilon^{-1}, \quad (2.5)$$

où le minimum est choisi parmi les  $K$  stabilisant  $G$  de manière interne.

Le Théorème 2.1 introduit un critère faisant intervenir des factorisations copremières normalisées de  $G$ . De par ses propriétés, c'est-à-dire

$$\begin{pmatrix} \widetilde{M} & \widetilde{N} \end{pmatrix} \begin{pmatrix} \widetilde{M} & \widetilde{N} \end{pmatrix}^* = I, \quad \left\| \begin{pmatrix} \widetilde{M} & \widetilde{N} \end{pmatrix} \right\|_\infty = 1,$$

nous pouvons effectuer la manipulation suivante :

$$\begin{aligned} \left\| \begin{pmatrix} I \\ K \end{pmatrix} (I - GK)^{-1} \widetilde{M}^{-1} \right\|_\infty &= \left\| \begin{pmatrix} I \\ K \end{pmatrix} (I - GK)^{-1} \widetilde{M}^{-1} \begin{pmatrix} \widetilde{M} & \widetilde{N} \end{pmatrix} \right\|_\infty, \\ &= \left\| \begin{pmatrix} I \\ K \end{pmatrix} (I - GK)^{-1} \begin{pmatrix} I & G \end{pmatrix} \right\|_\infty, \end{aligned}$$

qui fait disparaître les termes dûs à la factorisation copremière. En d'autres termes, (2.5) peut se réécrire en utilisant les transferts  $S$ ,  $GS$ ,  $KS$ ,  $KSG$  :

$$\left\| \begin{pmatrix} S & SG \\ KS & KSG \end{pmatrix} \right\|_\infty \leq \gamma \Leftrightarrow \|\Pi(K, G)\|_\infty \leq \gamma. \quad (2.6)$$

On remarque alors les quatre matrices de transferts introduites précédemment traduisent le compromis performance/robustesse. En effet, reprenant le schéma de la Figure 2.2, nous avons

$$\begin{pmatrix} y \\ -u \end{pmatrix} = \begin{pmatrix} S & SG \\ KS & KSG \end{pmatrix} \begin{pmatrix} \delta \\ d \end{pmatrix},$$

où  $y$  représente la mesure de la sortie du système,  $u$  la commande générée par le correcteur, et  $d$  et  $\delta$  les perturbations en entrée et sortie de système. Maximiser la grandeur de la famille de systèmes définie par leur incertitude sur les facteurs copremiers revient alors à minimiser la norme  $H_\infty$  d'un *problème 4 blocs* dont les transferts ont une signification physique bien particulière (voir la Section 2.2).

Afin de trouver un correcteur  $K$  stabilisant robustement  $G$ , nous préférons travailler sur la norme  $H_\infty$  du projecteur  $\Pi(K, G)$ . Nous définissons alors le problème de commande suivant qui sera central dans tout ce manuscrit.

**Définition 2.7. Problème de Commande Robuste (PCR)**

Étant donné  $\gamma > 1$ , trouver un correcteur  $K$  stabilisant  $G$  de manière interne et qui satisfasse :

$$\|\Pi(K, G)\|_\infty \leq \gamma.$$

Ce type de problème peut alors être traité à l'aide de techniques d'optimisation  $H_\infty$  classiques (algorithmes itératifs, voir [DF00, ACA<sup>+</sup>99]). Cependant, [GM88] montre qu'il existe une *solution explicite* au PCR, ce qui fait la force principale de cette approche. Cette solution se base sur le calcul de solutions définies positives des deux équations de Riccati algébriques suivantes :

$$\mathcal{R} := XA + A^T X - XBB^T X + C^T C = 0, \quad (2.7)$$

$$\widetilde{\mathcal{R}} := YA^T + AY - YC^T CY + BB^T = 0. \quad (2.8)$$

Ces deux équations ont les propriétés suivantes.

**Théorème 2.2.** [GM89, Th. A.1, Rem. A.2][ZDG96, Coroll. 13.8] Soit  $(A, B, C)$  une représentation d'état d'un système linéaire  $G$ . Si  $(A, B)$  est commandable et  $(C, A)$  observable, alors il existe une unique solution définie positive aux équations de Riccati algébriques  $\mathcal{R} = 0$  et  $\widetilde{\mathcal{R}} = 0$ .

La solution explicite du PCR est alors donné par le théorème suivant.

**Théorème 2.3.** [GM88] Soit  $X$  (resp.  $Y$ ) l'unique solution définie positive de  $\mathcal{R} = 0$  (resp.  $\tilde{\mathcal{R}} = 0$ ). La valeur minimale  $\gamma_{\text{opt}}$  de  $\gamma$  satisfaisant (2.5) est donnée par

$$\gamma_{\text{opt}} = \varepsilon_{\text{opt}}^{-1} = \left(1 - \left\| \begin{pmatrix} \tilde{N} \\ \tilde{M} \end{pmatrix} \right\|_H^2 \right)^{-1/2} = \sqrt{1 + \lambda_{\max}(YX)}, \quad (2.9)$$

où  $\lambda_{\max}(M)$  désigne la valeur réelle propre maximale d'une matrice  $M \in \mathbb{R}^{n \times n}$  et  $\|\bullet\|_H$  la norme de Hankel [ZDG96, Chap. 8].

**Remarque 2.6.**  $\gamma_{\text{opt}}$  peut être perçu comme une marge de stabilité. En effet, par définition (2.3),  $\varepsilon$  définit la "grandeur" de la famille de modèles stabilisables par un correcteur  $K$  donné.

Dans cette perspective, [Vin00] effectue le lien entre  $\gamma_{\text{opt}}$  et les marges de gain  $m_G$  et de phase  $m_\Phi$  assurées. Si  $M_G(K, G)$  désigne la marge de gain réelle de la boucle ouverte  $KG$ , et si  $M_\Phi(K, G)$  désigne sa marge de phase, alors nous avons :

$$\begin{cases} M_G(K, G) \geq m_G := \frac{1 + \gamma^{-1}}{1 - \gamma^{-1}}, \\ M_\Phi(K, G) \geq m_\Phi := 2 \arcsin(\gamma^{-1}). \end{cases} \quad (2.10)$$

### 2.3.3 Conception d'un correcteur

Maintenant que l'on connaît la valeur minimale de  $\gamma$  satisfaisant (2.5), nous pouvons déterminer à priori la plus grande boule de modèle stabilisable par un correcteur donné. Il reste cependant à exhiber un tel correcteur. [GM89] montre que les correcteurs minimisant  $\gamma$  peuvent être obtenus en résolvant un problème d'*extension de Nehari*. Ils concluent en donnant une paramétrisation de l'ensemble des correcteurs satisfaisant (2.5). Dans la suite, nous notons la représentation d'état  $(A, B, C, D)$  d'un système  $G$  par  $\left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$ .

**Théorème 2.4.** [GM89] Soit  $X$  (resp.  $Y$ ) la solution définie positive de  $\mathcal{R} = 0$  (resp.  $\tilde{\mathcal{R}} = 0$ ). Soient  $\gamma > 1$  satisfaisant (2.5),  $\zeta := (\gamma^2 - 1)^{1/2}$ ,  $Z_\gamma := (I_n + (XY - \gamma^2 I_n))^{-T} \in \mathbb{R}^{n \times n}$  et  $\Phi \in RH_\infty$  tel que  $\|\Phi\|_\infty \leq 1$ . De plus, soient les transferts suivants :

$$\begin{aligned} L_{1,1} &:= \left[ \begin{array}{c|c} A - B B^T X & -\gamma^2 Z_\gamma B \\ \hline -B^T X & I_m \end{array} \right] \in \mathbb{R}^{m \times m}, \\ L_{1,2} &:= \left[ \begin{array}{c|c} A - B B^T X & \gamma^2 \zeta^{-1} Z_\gamma Y C^T \\ \hline -B^T X & \mathbb{O}_{m \times p} \end{array} \right] \in \mathbb{R}^{m \times p}, \\ L_{2,1} &:= \left[ \begin{array}{c|c} A - B B^T X & -\gamma^2 Z_\gamma B \\ \hline C & \mathbb{O}_{p \times m} \end{array} \right] \in \mathbb{R}^{p \times m}, \\ L_{2,2} &:= \left[ \begin{array}{c|c} A - B B^T X & \gamma^2 \zeta^{-1} Z_\gamma Y C^T \\ \hline C & -\zeta^{-1} I_p \end{array} \right] \in \mathbb{R}^{p \times p}, \end{aligned} \quad L := \begin{bmatrix} L_{1,1} & L_{1,2} \\ L_{2,1} & L_{2,2} \end{bmatrix} \in \mathbb{R}^{(m+p) \times (m+p)}.$$

Les correcteurs satisfaisant (2.4), c'est-à-dire stabilisant tout  $G_\Delta$  dans  $\mathcal{G}_\varepsilon$  avec  $\varepsilon = \gamma^{-1}$ , sont donnés par :

$$K := (L_{1,1} \Phi + L_{1,2}) (L_{2,1} \Phi + L_{2,2})^{-1}.$$

Le choix d'un élément  $\Phi$  de la boule unité de  $RH_\infty$  étant libre à priori, il est d'usage de considérer le correcteur correspondant à  $\Phi = 0$ . Ce correcteur est appelé *correcteur central*.

**Corollaire 2.1.** Soit  $X$  (resp.  $Y$ ) la solution définie positive de  $\mathcal{R} = 0$  (resp.  $\tilde{\mathcal{R}} = 0$ ). De plus, soit  $\gamma > 1$  satisfaisant (2.5). Finalement, soient les matrices suivantes :

$$\begin{cases} Z_\gamma := (I_n + (XY - \gamma^2 I_n))^{-T}, \\ A_\gamma := A - B B^T X + \gamma^2 Z_\gamma Y C^T C, \\ B_\gamma := -\gamma^2 Z_\gamma Y C^T, \\ C_\gamma := B^T X. \end{cases}$$

Le *correcteur central* pour  $\gamma > \gamma_{\text{opt}}$  a alors pour représentation d'état  $(A_\gamma, B_\gamma, C_\gamma)$ . Sa fonction de transfert est donnée par :

$$K_c(s) := C_\gamma (s I_n - A_\gamma)^{-1} B_\gamma. \quad (2.11)$$

*Démonstration.* Le correcteur central est obtenu pour  $\Phi = 0$ . Alors, en reprenant les notations du Théorème 2.4, on a :

$$K_c = L_{1,2} L_{2,2}^{-1}.$$

Le schéma de la Figure 2.5 montre que le correcteur  $K_c$  peut se réécrire sous la forme d'un feedback. La nouvelle représentation de  $K_c$  indique que la sortie du correcteur  $y_1$  dépend de

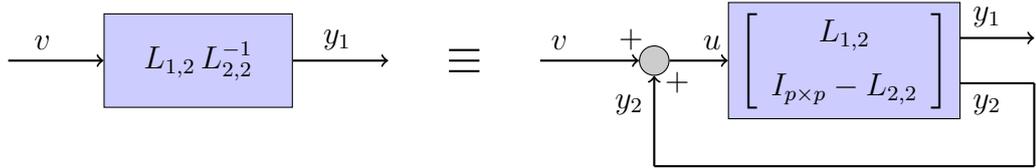


FIGURE 2.5 – Construction du correcteur central : équivalence de deux écritures

deux transferts  $L_{1,2}$  et  $I_{p \times p} - L_{2,2}$ . Ces deux transferts s'écrivent :

$$L_{1,2} = \left[ \begin{array}{c|c} A - B B^T X & \gamma^2 \zeta^{-1} Z_\gamma Y C^T \\ \hline -B^T X & \mathbb{O}_{m \times p} \end{array} \right] \in \mathbb{R}^{m \times p},$$

$$I_{p \times p} - L_{2,2} = \left[ \begin{array}{c|c} A - B B^T X & \gamma^2 \zeta^{-1} Z_\gamma Y C^T \\ \hline -C & I_p + \zeta^{-1} I_p \end{array} \right] \in \mathbb{R}^{p \times p}.$$

Remarquons que l'équation du vecteur d'état est la même pour ces deux transferts. Ainsi, leur vecteur d'état est le même. Notons  $x_L$  ce vecteur d'état commun. Nous avons alors les équations suivantes :

$$\begin{cases} \dot{x}_L = (A - B B^T X) x_L + \gamma^2 \zeta^{-1} Z_\gamma Y C^T u, \\ y_1 = -B^T X x_L, \\ y_2 = -C x_L + (I_p + \zeta^{-1} I_p) u, \\ u = y_2 + v. \end{cases}$$

Le résultat s'obtient alors par élimination successive des variables  $u$  et  $y_2$  (en combinant les équations précédentes) afin d'obtenir une équation d'état en  $x_L$  et  $v$ .  $\square$

**Remarque 2.7.** Pour un système  $G$  à paramètres fixés, le calcul du gamma optimal (2.9) et d'un correcteur central (2.11) peut se faire à l'aide, par exemple, de la commande `ncfsyn` de MATLAB. Celle-ci est valable dans le cas des systèmes MIMO avec prise en compte de transfert direct  $D$ .

### 2.3.4 Cas particulier des systèmes SISO

Supposons maintenant que  $G$  est un système SISO de fonction de transfert minimale  $G$  d'ordre  $n$ . Supposons que  $G$  est strictement propre (c'est-à-dire  $\lim_{|s| \rightarrow \infty} |G(s)| = 0$ ) :

$$G(s) = \frac{\sum_{k=0}^{n-1} c_k s^k}{s^n + \sum_{k=0}^{n-1} a_k s^k}.$$

Notons

$$N := \sum_{k=0}^{n-1} c_k s^k, \quad D := s^n + \sum_{k=0}^{n-1} a_k s^k,$$

et  $a := (a_0, \dots, a_{n-1})$  et  $c := (c_0, \dots, c_{n-1})$ . Il existe une infinité de représentations d'état possibles de  $G$ . Considérons une représentation sous forme *compagne de commande*, que l'on note  $(A, B, C)$ , de transfert direct  $D$  nul. Nous noterons les coefficients de  $(A, B, C)$  comme suit :

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix} \in \mathbb{R}[a]^{n \times n}, \quad B := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}[c]^{n \times 1},$$

$$C := (c_0 \quad c_1 \quad \dots \quad c_{n-1}) \in \mathbb{R}^{1 \times n}.$$

Il est alors possible de lier la solution positive de  $\mathcal{R} = 0$  avec celle de  $\tilde{\mathcal{R}} = 0$ . Pour cela, définissons la matrice  $\mathcal{P}$  suivante

$$\mathcal{P} := (\mathcal{P}_1 \quad \dots \quad \mathcal{P}_i \quad \dots \quad \mathcal{P}_n), \quad (2.13)$$

avec :

$$\left\{ \begin{array}{l} \mathcal{P}_n := C^T, \\ \mathcal{P}_{n-1} := (A^T + a_{n-1} I_n) C^T, \\ \vdots \\ \mathcal{P}_2 := \left( (A^T)^{n-2} + a_{n-1} (A^T)^{n-3} + \dots + a_2 I_n \right) C^T, \\ \mathcal{P}_1 := \left( (A^T)^{n-1} + a_{n-1} (A^T)^{n-2} + \dots + a_1 I_n \right) C^T. \end{array} \right.$$

Cette matrice permet le passage d'une forme canonique de commande à une forme compagne d'observation (voir Annexe A.5). Elle est aussi liée à la condition d'observabilité de  $G$  par la relation suivante.

**Proposition 2.1.** Supposons que  $G$  n'est pas observable. Alors,  $\mathcal{P}$  n'est pas inversible.

*Démonstration.* La matrice d'observabilité de Kalman est définie par :

$$\mathcal{O} := \begin{pmatrix} C \\ C A \\ \vdots \\ C A^{n-2} \\ C A^{n-1} \end{pmatrix} \in \mathbb{R}[a, c]^{n \times n}. \quad (2.14)$$

Nous notons son déterminant :

$$\Delta_o(a, c) := \det(\mathcal{O}). \quad (2.15)$$

Lorsque  $\Delta_o = 0$ , alors le système  $G$  n'est pas observable (chute du rang de  $\mathcal{O}$ ). La variété  $\Delta_o(a, c) = 0$  représente le lieu des paramètres  $a$  et  $c$  à éviter pour trouver une solution au PCR. Remarquons alors que :

$$\underbrace{\begin{pmatrix} C (A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I_n) \\ C (A^{n-2} + a_{n-1} A^{n-3} + \cdots + a_2 I_n) \\ \vdots \\ C (A + a_{n-1} I_n) \\ C \end{pmatrix}}_{\mathcal{P}^T} = \underbrace{\begin{pmatrix} 1 & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \\ 0 & 1 & a_{n-1} & \cdots & a_3 & a_2 \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1} & \vdots \\ 0 & 0 & \cdots & 0 & 1 & a_{n-1} \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}}_{T_n} \times \underbrace{\begin{pmatrix} C A^{n-1} \\ C A^{n-2} \\ \vdots \\ C A \\ C \end{pmatrix}}_{\mathcal{O}_i}.$$

La permutation de deux lignes d'un déterminant multipliant ce dernier par  $-1$ , nous en déduisons :

$$\det(\mathcal{P}^T) = \det(\mathcal{P}) = \det(T_n) \times \det(\mathcal{O}_i) = \begin{cases} \det(T_n) \times (-1)^{\frac{n}{2}} \Delta_o, & \text{si } n \text{ pair,} \\ \det(T_n) \times (-1)^{\frac{n-1}{2}} \Delta_o, & \text{si } n \text{ impair.} \end{cases}$$

Pour  $q \in \mathbb{Q}$ , notons  $\text{rd}(q)$  l'arrondi inférieur de  $q$  dans  $\mathbb{N}$ . Nous avons :

$$\det(\mathcal{P}) = (-1)^{\text{rd}(n/2)} \Delta_o. \quad (2.16)$$

Cette dernière égalité montre que  $\mathcal{P}$  n'est pas inversible si  $\Delta_o = 0$ , c'est-à-dire quand le système n'est pas observable.  $\square$

Cette condition de non observabilité  $\Delta_o(a, c) = 0$  se retrouvera dans l'intégralité de ce mémoire.

Disposant d'une telle matrice  $\mathcal{P}$ , nous avons le théorème suivant.

**Théorème 2.5.** Soit  $G$  un système linéaire SISO d'ordre  $n$ , strictement propre, de représentation d'état sous forme canonique de commande  $(A, B, C)$ . Supposons que  $G$  est commandable et observable. Soit  $X$  l'unique solution définie positive de  $\mathcal{R} = 0$ . Soient  $\mathcal{P}$  la matrice définie par (2.13) et  $\mathcal{Q} := \mathcal{P}^{-1}$ . Alors,  $Y := \mathcal{Q} X \mathcal{Q}$  est l'unique solution définie positive de  $\tilde{\mathcal{R}} = 0$ .

*Démonstration.* Considérons une solution définie positive de  $\mathcal{R} = 0$ . Nous avons :

$$X A + A^T X - X B B^T X + C^T C = 0.$$

Multiplications cette dernière équation à droite et à gauche par  $\mathcal{Q}$  :

$$\mathcal{Q} X A \mathcal{Q} + \mathcal{Q} A^T X \mathcal{Q} - \mathcal{Q} X B B^T X \mathcal{Q} + \mathcal{Q} C^T C \mathcal{Q} = 0. \quad (2.17)$$

Nous rappelons les propriétés suivantes de  $\mathcal{Q}$  (voir Annexe A.5) :

$$\begin{cases} A = \mathcal{Q} A^T \mathcal{Q}^{-1}, \\ B = \mathcal{Q} C^T, \\ C = B^T \mathcal{Q}^{-1}. \end{cases} \quad (2.18)$$

En remplaçant les expressions de  $C$  et  $B$  données dans (2.18) dans (2.17), nous obtenons :

$$(\mathcal{Q} X \mathcal{Q}) A^T + A (\mathcal{Q} X \mathcal{Q}) - (\mathcal{Q} X \mathcal{Q}) C^T C (\mathcal{Q} X \mathcal{Q}) + B B^T = 0.$$

Soit  $\mathcal{Y} := \mathcal{Q} X \mathcal{Q}$ . Il apparaît alors que  $\mathcal{Y}$  est solution de  $\tilde{\mathcal{R}} = 0$ . De plus,  $\mathcal{Q}$  est une matrice symétrique inversible, ce qui induit :

$$\begin{aligned} \det(\mathcal{Y} - \lambda I_n) &= \det(\mathcal{Q} X \mathcal{Q} - \lambda I_n) = \det(\mathcal{Q}^{-1}) \det(\mathcal{Q} X \mathcal{Q} - \lambda I_n) \det(\mathcal{Q}), \\ &= \det(X \mathcal{Q} \mathcal{Q} - \lambda I_n), \\ &= \det(X \mathcal{Q}^T \mathcal{Q} - \lambda I_n). \end{aligned}$$

Par ailleurs,  $\mathcal{Q}^2 = \mathcal{Q}^T \mathcal{Q}$  est une matrice définie positive (voir Annexe C). L'expression précédente fait alors apparaître le produit de deux matrices définies positives  $X$  et  $\mathcal{Q}^T \mathcal{Q}$ . Sachant que les valeurs propres de ce produit sont strictement positives (voir Annexe C), les valeurs propres de  $\mathcal{Y}$  le sont aussi.

Enfin, sachant que  $\mathcal{Q}$  et  $X$  sont symétriques et que  $\mathcal{Y} = \mathcal{Q} X \mathcal{Q}$ , alors  $\mathcal{Y} = \mathcal{Y}^T$ . En d'autres termes,  $\mathcal{Y}$  est symétrique, a ses valeurs propres réelles strictement positives : elle est donc définie positive. Ainsi,  $Y := \mathcal{Y}$  est solution de  $\tilde{\mathcal{R}} = 0$ .  $\square$

Ce résultat est très intéressant en pratique puisque nous pouvons réduire la résolution du PCR (dans le cas des systèmes SISO) à la résolution d'une *unique équation de Riccati algébrique*  $\mathcal{R} = 0$ . En effet, en notant  $\lambda_*(M)$  la valeur *absolue maximale des valeurs propres réelles* d'une matrice  $M$ , nous avons le théorème suivant.

**Théorème 2.6.** Soit  $G$  un système SISO, strictement propre, de représentation d'état sous forme canonique de commande  $(A, B, C)$ . Soit l'équation de Riccati suivante :

$$\mathcal{R} := X A + A^T X - X B B^T X + C^T C = 0.$$

Soit  $X$  l'unique solution définie positive de  $\mathcal{R} = 0$ . La valeur minimale de  $\gamma$  satisfaisant (2.5) est donnée par :

$$\gamma_{\text{opt}} = \varepsilon_{\text{opt}}^{-1} = \sqrt{1 + \lambda_{\max}(Y X)} = \sqrt{1 + (\lambda_*(\mathcal{Q} X))^2}. \quad (2.19)$$

*Démonstration.*  $X$  et  $Y := \mathcal{Q} X \mathcal{Q}$  sont deux matrices définies positives. La Proposition C.2 de l'Annexe C permet alors d'établir directement que les valeurs propres de  $Y X$  sont réelles strictement positives. De plus, nous avons  $Y X = (\mathcal{Q} X)^2$ . Par conséquent, si  $\lambda$  est une valeur propre de  $\mathcal{Q} X$ , alors  $\lambda^2$  est une valeur propre de  $Y X$ . Or,  $\lambda^2$  est un réel strictement positif, donc  $\lambda$  est réel. Les valeurs propres de  $\mathcal{Q} X$  sont donc nécessairement toutes réelles. De plus, nous avons alors :

$$\det(Y X - \lambda^2 I_n) = \det(\mathcal{Q} X - \lambda I_n) \det(\mathcal{Q} X + \lambda I_n). \quad (2.20)$$

Ainsi, d'après (2.20),  $\lambda^2$  est valeur propre réelle de  $Y X$  si et seulement si  $\lambda$  ou  $-\lambda$  est valeur propre de  $\mathcal{Q} X$ . En particulier,  $\lambda^2$  est valeur propre réelle maximale de  $Y X$  si et seulement si  $\lambda$  ou  $-\lambda$  est valeur propre réelle maximale de  $\mathcal{Q} X$ . En d'autres termes, nous avons :

$$\lambda_{\max}(Y X) = (\lambda_*(\mathcal{Q} X))^2.$$

De plus, le Théorème 2.3 montre que

$$\gamma_{\text{opt}} = \sqrt{1 + \lambda_{\max}(Y X)} = \sqrt{1 + \lambda_{\max}((\mathcal{Q} X)^2)},$$

et nous déduisons :

$$\gamma_{\text{opt}} = \sqrt{1 + (\lambda_*(\mathcal{Q} X))^2}.$$

$\square$

## 2.4 Lien avec la commande Linéaire Quadratique Gaussienne

Rappelons que, pour un système  $G$  de représentation d'état  $(A, B, C)$  et disposant de  $\gamma$  majorant  $\gamma_{\text{opt}}$  défini par (2.9), les matrices de la représentation d'état  $(A_\gamma, B_\gamma, C_\gamma)$  du correcteur central  $K_c$  obtenu au Corollaire 2.1 peuvent s'écrire

$$\begin{cases} B_\gamma := \gamma^2 ((\gamma^2 - 1) I - Y X)^{-1} Y C^T, \\ A_\gamma := -B B^T X + A - B_\gamma C, \\ C_\gamma := B^T X, \end{cases}$$

et nous avons :

$$\begin{cases} \dot{x}_k = A_\gamma x_k + B_\gamma y, \\ u = C_\gamma x_k. \end{cases}$$

Lorsque  $\gamma$  tend vers l'infini, nous obtenons les matrices suivantes

$$\begin{cases} B_\infty := \lim_{\gamma \rightarrow \infty} B_\gamma = \lim_{\gamma \rightarrow \infty} \frac{\gamma^2}{\gamma^2 - 1} (I - \frac{1}{\gamma^2 - 1} Y X)^{-1} Y C^T = Y C^T, \\ A_\infty := \lim_{\gamma \rightarrow \infty} A_\gamma = A - B B^T X - Y C^T C, \\ C_\infty := \lim_{\gamma \rightarrow \infty} C_\gamma = B^T X, \end{cases}$$

c'est-à-dire que le correcteur s'écrit

$$\begin{cases} \dot{x}_k = (A - B B^T X - Y C^T C) x_k + Y C^T y, \\ u = B^T X x_k, \end{cases}$$

où  $X$  et  $Y$  sont les solutions définies positives des équations de Riccati suivantes :

$$\begin{aligned} \mathcal{R} &:= X A + A^T X - X B B^T X + C^T C = 0, \\ \tilde{\mathcal{R}} &:= Y A^T + A Y - Y C^T C Y + B B^T = 0. \end{aligned}$$

Le correcteur obtenu correspond alors à un correcteur obtenu par synthèse Linéaire Quadratique Gaussienne (dite LQG) [ACA<sup>+</sup>99]. Celle-ci consiste à la conception d'un correcteur par retour d'état minimisant un critère quadratique (dit LQ), tout en estimant l'état via un *filtre de Kalman*.

Plus particulièrement, l'équation  $\mathcal{R} = 0$  correspond alors à l'étape de conception du correcteur par retour d'état  $u = -B^T X x$  qui minimise le critère LQ suivant

$$\int_0^\infty (y^T y + u^T u) dt,$$

où  $u$  désigne la commande et  $y = C x$  la mesure.

L'autre équation  $\tilde{\mathcal{R}} = 0$  correspond à l'estimation de l'état  $x$  du système, estimé noté par  $\hat{x}$ , par un filtre de Kalman dans le cas où le bruit Gaussien sur la commande  $u$  et sur la mesure  $y$  sont unitaires :

$$\hat{\dot{x}} = A \hat{x} - B u + Y C^T (y - C \hat{x}).$$

Par conséquent, la résolution du problème de commande robuste  $H_\infty$  introduit dans ce chapitre implique la résolution du problème de commande LQG. Il suffit pour cela de faire tendre  $\gamma$  vers l'infini. Dans ce cas, nous pouvons nous passer de la partie recherche de valeurs propres de  $YX$  (ou  $\mathcal{L}X$ ).

## 2.5 Commande $H_\infty$ par Loop-Shaping

Nous disposons à présent d'outils permettant de stabiliser un modèle tout en garantissant une marge de stabilité définie via les facteurs copremiers de celui-ci. De plus, la Section 2.2 a formalisé des objectifs fréquentiels à atteindre sur quatre fonctions de transfert. La référence [MG92] propose une procédure utilisant les outils  $H_\infty$  présentés en Section 2.3 et sur le concept de *Loop-Shaping* ou *modelage de la boucle ouverte*. Dans un souci de simplicité, nous exposons cette procédure dans le cas des systèmes SISO, mais celle-ci s'étend aux systèmes MIMO par des considérations de valeurs singulières.

### 2.5.1 Modelage de la boucle ouverte

Nous énonçons en introduction que le concept du Loop-Shaping permettait d'atteindre des objectifs fréquentiels sur les quatre fonctions  $S$ ,  $GS$ ,  $KS$  et  $KGS$  via l'ajout de pondérations fréquentielles. Ce concept, introduit dans [MG92] (et dont on peut trouver une synthèse en langue française dans [DF00] par exemple), se base dans un premier temps sur les remarques suivantes.

**Remarque 2.8.** Soit un système  $P$  et un correcteur  $C$  stabilisant  $P$  de manière interne.

- Supposons que  $|PC| \gg 1$ . Alors, nous avons :

$$\left| \frac{1}{1+PC} \right| \approx \left| \frac{1}{PC} \right|, \quad (2.21)$$

$$\left| \frac{P}{1+PC} \right| \approx \left| \frac{1}{C} \right|. \quad (2.22)$$

En pratique, cette supposition est valable en basse fréquence (présence d'un ou deux intégrateurs purs dans nos systèmes mécaniques).

- Supposons que  $|PC| \ll 1$ . Alors, nous avons :

$$\left| \frac{C}{1+PC} \right| \approx |C|, \quad (2.23)$$

$$\left| \frac{PC}{1+PC} \right| \approx |PC|. \quad (2.24)$$

En pratique, cette supposition est valable en haute fréquence (présence d'intégrateurs purs, et de filtrage haute fréquence lié au capteur).

Les remarques précédentes montrent qu'une manipulation du gain et un modelage de la boucle ouverte permettent d'attendre les objectifs fixés en Section 2.2 sur les fonctions de sensibilité. Ceci étant dit, il reste à énoncer une méthode permettant de modeler la dite boucle ouverte. Cette méthode se base sur la propriété suivante.

**Proposition 2.2.** Soit  $P$  un système SISO. Soit  $W \in \mathbb{R}(s)$  une pondération fréquentielle telle que  $WP$  ne contienne pas de simplification pôle/zéro. Définissons  $G := WP$ , que l'on appelle *système fictif*. Soit un correcteur  $K$  stabilisant  $G$  de manière interne. Alors le correcteur  $C := WK$  stabilise  $P$  de manière interne.

*Démonstration.* Considérons une factorisation copremière de  $P$  :

$$P = \frac{N}{M}, \quad (M, N) \in RH_\infty^2.$$

De même, considérons une factorisation copremière de  $W$  :

$$W = \frac{A}{B}, \quad (A, B) \in RH_\infty^2.$$

Sachant qu'il n'existe pas de simplification pôle/zéro dans  $G := WP$ , alors les éléments de  $RH_\infty$  suivants

$$\begin{cases} N_0 := AN \in RH_\infty, \\ M_0 := BM \in RH_\infty, \end{cases}$$

définissent une factorisation copremière de  $G$  [Vid85] :

$$G = WP = \frac{AN}{BM} = \frac{N_0}{M_0}.$$

Soit  $K$  un correcteur stabilisant  $G$  de manière interne. Considérons une factorisation copremière de  $K$  :

$$K = \frac{V_0}{U_0}, \quad (U_0, V_0) \in RH_\infty^2.$$

Soit  $Z := M_0 U_0 + N_0 V_0$ . Alors, nous avons :

$$\Pi(K, G) = \frac{1}{Z} \begin{pmatrix} M_0 U_0 & N_0 U_0 \\ M_0 V_0 & N_0 V_0 \end{pmatrix}.$$

Comme le produit ou la somme de fonctions de  $RH_\infty$  est dans  $RH_\infty$ , il vient  $Z \in RH_\infty$ ,  $M_0 U_0 \in RH_\infty$ ,  $N_0 U_0 \in RH_\infty$ ,  $M_0 V_0 \in RH_\infty$  et  $N_0 V_0 \in RH_\infty$ .

Montrons maintenant que :

$$\Pi(K, G) \in RH_\infty^{2 \times 2} \quad \Leftrightarrow \quad \frac{1}{Z} \in RH_\infty.$$

• Supposons que  $\Pi(K, G) \in RH_\infty^{2 \times 2}$ . Comme le couple  $(U_0, V_0) \in RH_\infty^2$  satisfait une équation de Bézout [ZDG96, Chap. 5.4]

$$U_0 \alpha_K + V_0 \beta_K = 1,$$

où  $(\alpha_K, \beta_K) \in RH_\infty^2$ , nous avons :

$$\begin{pmatrix} \alpha_K & \beta_K \end{pmatrix} \Pi(K, G) = \frac{1}{Z} \begin{pmatrix} M_0 & N_0 \end{pmatrix} \in RH_\infty^{1 \times 2}.$$

Nous avons alors  $\frac{M_0}{Z} \in RH_\infty$  et  $\frac{N_0}{Z} \in RH_\infty$ . Considérons alors l'équation de Bézout associée à la factorisation copremière de  $G$  :

$$\exists(\alpha_G, \beta_G) \in RH_\infty^2, \quad M_0 \alpha_G + N_0 \beta_G = 1.$$

De l'équation précédente, en utilisant que  $\alpha_G \in RH_\infty$ ,  $\beta_G \in RH_\infty$ ,  $\frac{M_0}{Z} \in RH_\infty$  et  $\frac{N_0}{Z} \in RH_\infty$ , alors nous déduisons :

$$\frac{1}{Z} = \frac{M_0}{Z} \alpha_G + \frac{N_0}{Z} \beta_G \in RH_\infty.$$

Nous avons montré que si  $\Pi(K, G) \in RH_\infty^{2 \times 2}$  alors  $\frac{1}{Z} \in RH_\infty$ .

• L'implication inverse est triviale.

Soit  $C := W K = \frac{V_0 A}{U_0 B}$ . Alors, nous avons :

$$\Pi(C, P) = \begin{pmatrix} \frac{1}{1+CP} & \frac{P}{1+CP} \\ \frac{C}{1+CP} & \frac{CP}{1+CP} \end{pmatrix} = \begin{pmatrix} \frac{1}{1+KG} & \frac{GW^{-1}}{1+KG} \\ \frac{KW}{1+KG} & \frac{KG}{1+KG} \end{pmatrix} = \frac{1}{Z} \begin{pmatrix} M_0 U_0 & N_0 U_0 B A^{-1} \\ M_0 V_0 A B^{-1} & N_0 V_0 \end{pmatrix}.$$

En notant que  $N_0 = A N$  et  $M_0 = B M$ , nous pouvons réécrire  $\Pi(C, P)$  de la manière suivante :

$$\Pi(C, P) = \frac{1}{Z} \begin{pmatrix} M_0 U_0 & U_0 B N \\ M V_0 A & N_0 V_0 \end{pmatrix}.$$

Nous avons  $M_0 U_0 \in RH_\infty$ ,  $U_0 B N \in RH_\infty$ ,  $M V_0 A \in RH_\infty$ ,  $N_0 V_0 \in RH_\infty$  car ces éléments sont des produits d'éléments de  $RH_\infty$ . De plus,  $1/Z \in RH_\infty$  donc  $\Pi(C, P) \in RH_\infty^{2 \times 2}$ .

Ainsi, notant  $U = B U_0 \in RH_\infty$  et  $V = A V_0 \in RH_\infty$ , nous déduisons que le correcteur

$$C := \frac{U}{V} = W K$$

stabilise  $P$  de manière interne. □

**Remarque 2.9.**  $W$  peut être instable, notamment contenir des intégrateurs [MG92].

La pondération  $W$  peut être vue comme un moyen de modéliser à la fois la boucle ouverte et le correcteur. Les contraintes à imposer sur les quatre fonctions de transfert vues précédemment imposent des choix en haute et basse fréquences sur la pondération  $W$  :

- $|S| \ll 1$  en basse fréquence pour avoir une faible erreur sur la mesure lorsqu'une perturbation agit en sortie de modèle. Ceci peut se faire en imposant  $|L| \gg 1$  en basse fréquence. On pourra alors introduire des comportements type intégrateur dans  $W$ .
- $|LS| \ll 1$  en haute fréquence pour être robuste aux erreurs de modélisation. Ceci peut se faire en imposant  $|L| \ll 1$  en haute fréquence. On pourra alors introduire des filtres haute-fréquence dans  $W$ .
- $|GS|$  le plus petit possible pour rejeter les perturbations en entrée de système. Ceci peut se faire en imposant  $|L| \gg 1$  en basse fréquence. Là encore, on pourra introduire des comportements intégrateurs dans  $W$  en basse fréquence pour satisfaire ce critère.
- $|KS|$  ne doit pas être excessivement fort en haute fréquence pour limiter l'amplification des bruits et les saturations de l'actionneur. Ceci peut se faire en ajoutant des actions filtrantes dans  $K$  en haute fréquence, toujours au moyen de la pondération  $W$ .

Le gabarit grossier de la boucle ouverte aura alors l'aspect décrit dans la Figure 2.6.

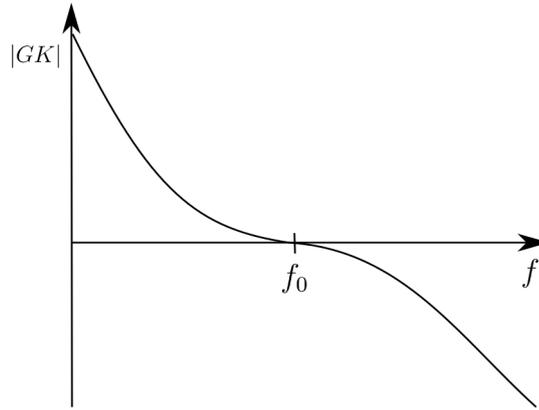


FIGURE 2.6 – Modélage de la boucle ouverte

Ainsi, il faut utiliser une pondération  $W$  comportant au moins une action intégrale en basse fréquence et un effet filtrant en haute fréquence. On règlera aussi la fréquence de passage à 0 dB de la boucle ouverte en ajustant le gain de  $W$ .

Ces techniques de gabarit sur la boucle ouverte ne dépendent pas d'une méthode de conception de correcteur particulière. Cependant, les références [GM89, MG92] motivent l'utilisation d'une méthode de commande minimisant un critère  $H_\infty$  en complément du Loop-Shaping.

### 2.5.2 Une procédure de synthèse de correcteur $H_\infty$ par Loop-Shaping

Soit  $P$  un système nominal SISO à asservir. La procédure énoncée dans [MG92] est la suivante :

1. Soit une pondération fréquentielle  $W$  propre non nécessairement stable. Ajuster cette pondération selon les recommandations de la Section 2.5.1. Créer le *système fictif* :

$$G := W P.$$

S'assurer que  $G$  ne contient pas de simplifications pôle/zéro, sinon revoir la pondération  $W$ . Considérer une représentation d'état  $(A, B, C, D)$  de  $G$ . Calculer  $\gamma_{\text{opt}}$  grâce à (2.9) pour déduire la solution optimale du PCR donné au Corollaire 2.1.

2. Si  $\gamma$  est grand, revenir à l'étape 1 en ajustant la pondération. Sinon, synthétiser un correcteur  $K$  stabilisant  $G$  et satisfaisant le PCR en utilisant le Corollaire 2.1.
3. Basculer la pondération dans le correcteur en utilisant le principe décrit dans la Proposition 2.2 :

$$C := W K.$$

Notons que la partie Loop-Shaping est utilisée pour modéliser la boucle ouverte à basse et haute fréquences, alors que la partie  $H_\infty$  est utilisée pour stabiliser de manière robuste le système fictif  $G$ . De plus, une telle méthode donne un *indicateur* ( $\gamma_{\text{opt}}$ ) sur la cohérence entre la pondération  $W$  utilisée, la phase du système physique  $G$  et la stabilité robuste du système bouclé. En effet, rappelons que  $\gamma_{\text{opt}}$  mesure la marge de stabilité maximale qu'un système peut atteindre. Ainsi, pour des  $\gamma_{\text{opt}}$  trop grands, il y a incohérence entre les spécifications de robustesse introduites dans le modelage de la boucle ouverte via  $W$  et celles réellement atteignables. En pratique, les valeurs de  $\gamma_{\text{opt}}$  utilisées varient entre 2 et 3.

## 2.6 Exemple d'application

Considérons la modélisation la plus fine des mesures du système présentée en Annexe G, à savoir

$$P_9 := \frac{P_0}{s} H_{m_1} H_{m_2} F_2 \left( \frac{1 - \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}}{1 + \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}} \right),$$

où :

$$H_{m_1} := \frac{s^2 + 2 \xi_0 \omega_0 s + \omega_0^2}{s^2 + 2 \xi_1 \omega_1 s + \omega_1^2}, \quad H_{m_2} := \frac{s^2 + 2 \xi_2 \omega_2 s + \omega_2^2}{s^2 + 2 \xi_3 \omega_3 s + \omega_3^2}, \quad F_2 := \frac{3.0429}{3.0430 + 3.3260 s + s^2}.$$

Après plusieurs démontages et remontages du viseur, on constate que les modes mécaniques se déplacent, tout en ayant toujours la même forme, à savoir une anti-résonance de fréquence plus faible que la résonance et des amortissements similaires. Fixons alors le cahier des charges suivant :

- Les amortissements  $\xi_i$  des modes mécaniques restent inchangés.
- La fréquence des modes varie de 20%.
- L'écart fréquentiel entre résonance et anti-résonance reste inchangé.

Considérons alors que les fréquences des modes prennent les valeurs suivantes

$$\begin{aligned} 0.8 \Omega_0 < \omega_0 < 1.2 \Omega_0, \quad \omega_1 &= \omega_0 + \Omega_1 - \Omega_0 \\ 0.8 \Omega_2 < \omega_2 < 1.2 \Omega_2, \quad \omega_3 &= \omega_2 + \Omega_3 - \Omega_2 \\ \Omega_0 = 1, \quad \Omega_1 = 1.03, \quad \Omega_2 = 1.74, \quad \Omega_3 = 1.75, \end{aligned} \quad (2.25)$$

et que les autres paramètres sont fixés aux valeurs suivantes :

$$\begin{aligned} P_0 = 6.65 \cdot 10^{-3}, \quad T_9 = 9.21 \cdot 10^{-1}, \\ \xi_0 = 1.08 \cdot 10^{-2}, \quad \xi_1 = 9.7 \cdot 10^{-2}, \quad \xi_2 = 5.74 \cdot 10^{-3}, \quad \xi_3 = 3.80 \cdot 10^{-3}. \end{aligned}$$

Nous notons par  $\mathcal{B}$  la boule de modèle suivante

$$\mathcal{B} := \{P_9 \mid 0.8 < \omega_0 < 1.2, \quad \omega_1 = \omega_0 + 0.03, \quad 1.39 < \omega_2 < 2.09, \quad \omega_3 = \omega_2 + 0.01\},$$

et par  $P_{9,n} \in \mathcal{B}$  le modèle nominal satisfaisant :

$$\forall i \in \{1, 2, 3, 4\}, \quad \omega_i = \Omega_i.$$

Cherchons à établir un asservissement robuste à ces variations paramétriques en reprenant la procédure de Loop-Shaping décrite précédemment. Considérons un échantillon de modèles dans  $\mathcal{B}$ . Nous choisissons ces modèles de telle manière à atteindre les bornes spécifiées en (2.25). Dans la Figure 2.7, les diagrammes de Bode de ces modèles tests sont représentés.

Afin de satisfaire les contraintes de robustesse, il faut que l'on crée un correcteur tel que l'inégalité sur la sensibilité complémentaire (2.2) soit vérifiée. Pour cela, dans la pondération  $W$ , on utilise un filtre passe-bas dont on affine la fréquence afin d'englober tous les modèles. On rajoute ensuite des effets intégrateurs en basse fréquence dans la pondération  $W$  pour rejeter les couples perturbateurs, puis enfin on ajuste le gain de la pondération pour régler la bande passante à 0dB de la boucle ouverte. Ce réglage demande un certain nombre d'itérations par essai-erreur. A chaque essai, le critère  $\gamma_{\text{opt}}$  appliqué au système  $G := WP$  est un indicateur de la qualité de la conception de la pondération. Si  $\gamma_{\text{opt}}$  est grand, il faut revoir la pondération, tandis qu'un  $\gamma_{\text{opt}}$  faible indique une bonne robustesse. On essaie d'ajuster le gain de la pondération afin d'avoir  $\gamma_{\text{opt}}$  autour de 3, ce qui garantit certaines marges de stabilité confortables en pratique (voir (2.10)). Finalement, on peut retenir la pondération suivante

$$W = \frac{15.4 (s + 0.1)^2}{s^2 (s + 0.7)},$$

qui est représentée dans la Figure 2.8. Pour une telle pondération, le critère  $H_\infty$  défini par (2.9) vaut alors

$$\gamma_{\text{opt}} \approx 2.97,$$

ce qui garantit des marges de stabilité confortables. Le correcteur correspondant stabilisant le système fictif  $G$  est donné par :

$$K = \frac{2.80(s + 0.702)(s^2 + 0.08189s + 0.002649)(s^2 + 3.326s + 3.042)}{(s + 2.911)(s^2 + 0.1951s + 0.009619)(s^2 + 1.574s + 1.603)} \\ \times \frac{(s^2 + 0.0199s + 1.062)(s^2 + 0.0133s + 3.071)(s^2 + 6.517s + 14.16)}{(s^2 + 0.03027s + 1.056)(s^2 + 0.01368s + 3.073)(s^2 + 6.33s + 14.1)}.$$

Le correcteur  $C = WK$  stabilisant la boule de modèle obtenu a alors la forme suivante (Figure 2.9) :

$$C = \frac{43.19(s + 0.702)(s + 0.1)^2(s^2 + 0.08189s + 0.002649)(s^2 + 3.326s + 3.042)}{s^2(s + 0.7)(s + 2.911)(s^2 + 0.1951s + 0.009619)(s^2 + 1.574s + 1.603)} \\ \times \frac{(s^2 + 0.0199s + 1.062)(s^2 + 0.0133s + 3.071)(s^2 + 6.517s + 14.16)}{(s^2 + 0.03027s + 1.056)(s^2 + 0.01368s + 3.073)(s^2 + 6.33s + 14.1)}.$$

Nous pouvons alors vérifier que l'ensemble de nos modèles est stable, soit en vérifiant que l'inégalité (2.2) est vérifiée pour chaque modèle échantillon (Figure 2.10), soit en analysant les diagrammes de Black de chaque boucle ouverte pour chaque modèle échantillon (Figure 2.11).

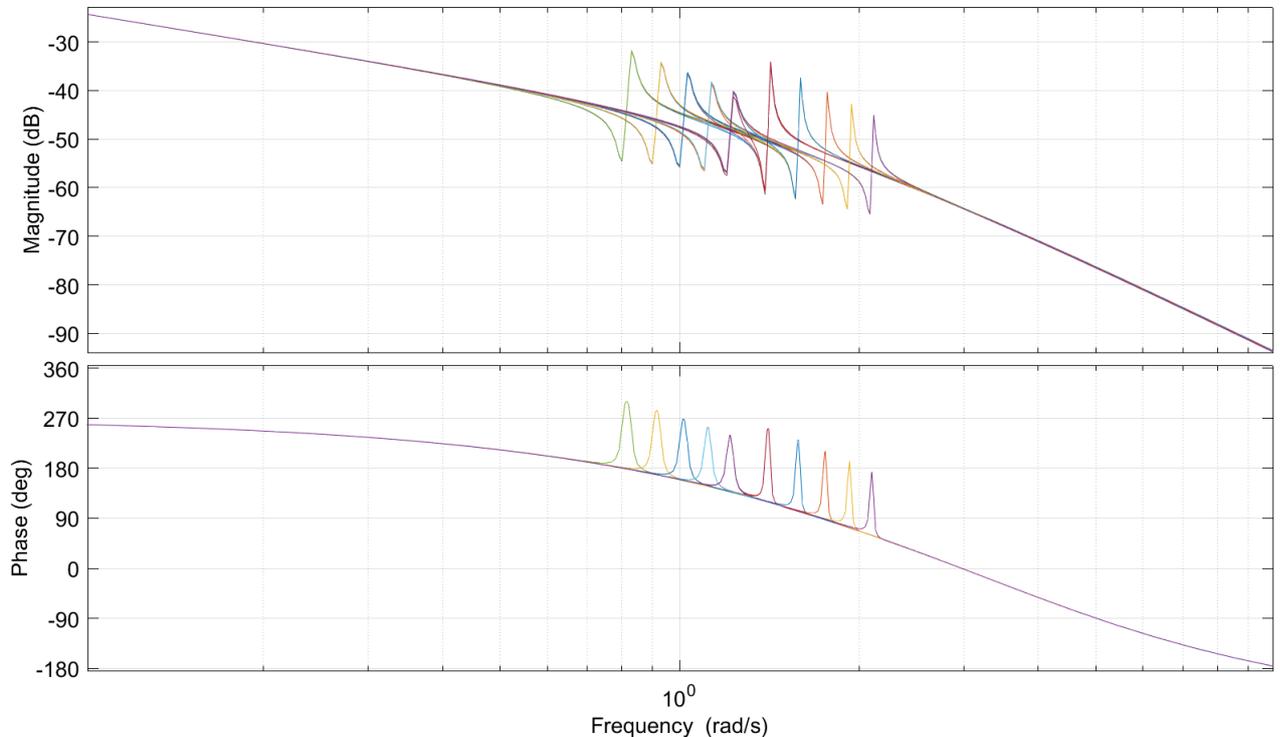


FIGURE 2.7 – Diagramme de Bode des échantillons de modèles  $P$  choisis dans  $\mathcal{B}$

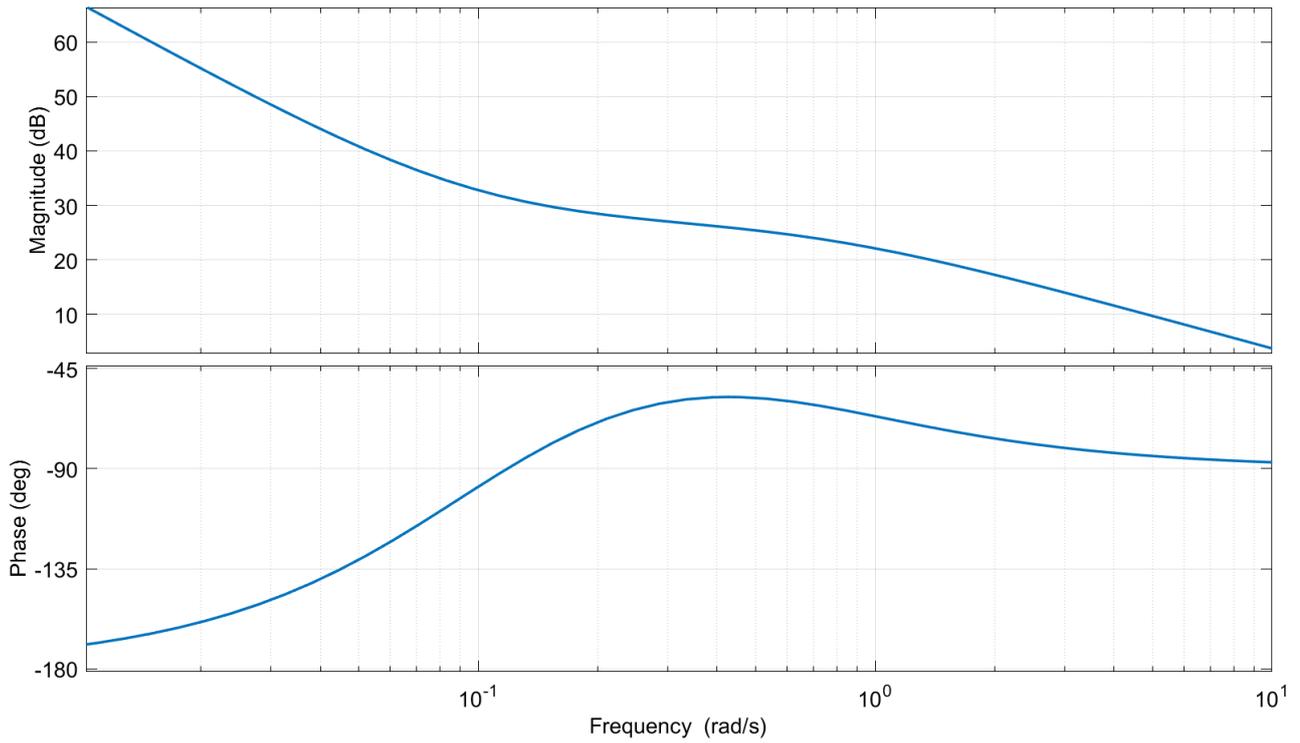


FIGURE 2.8 – Diagramme de Bode de la pondération Loop-Shaping utilisée

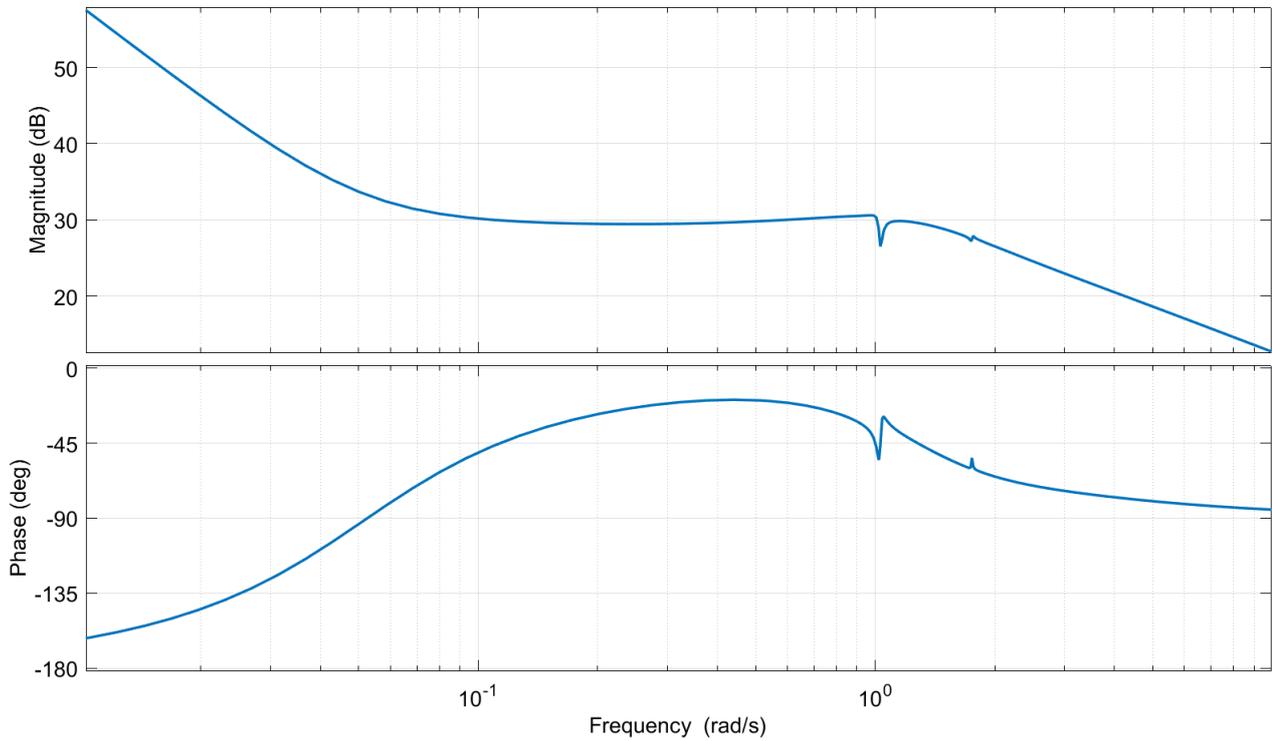


FIGURE 2.9 – Diagramme de Bode du correcteur  $C$  obtenu

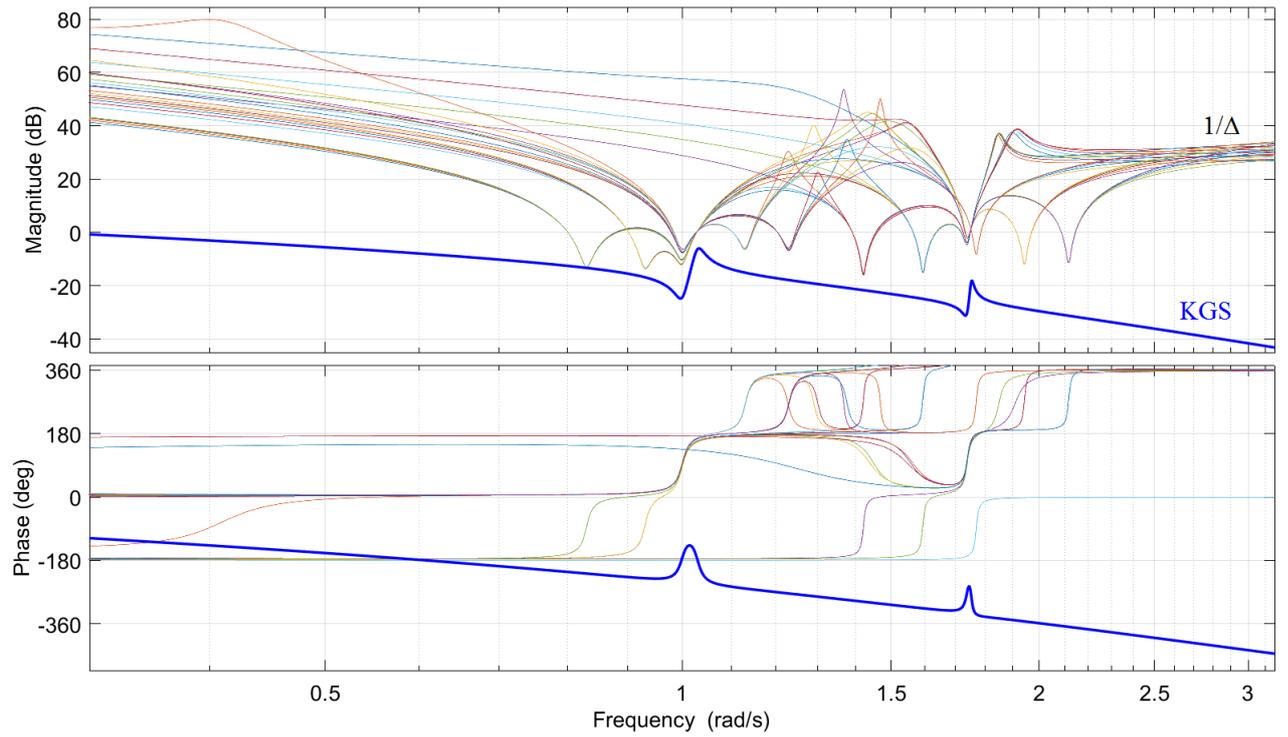


FIGURE 2.10 – Prise en compte d'incertitudes de modèle via la sensibilité complémentaire

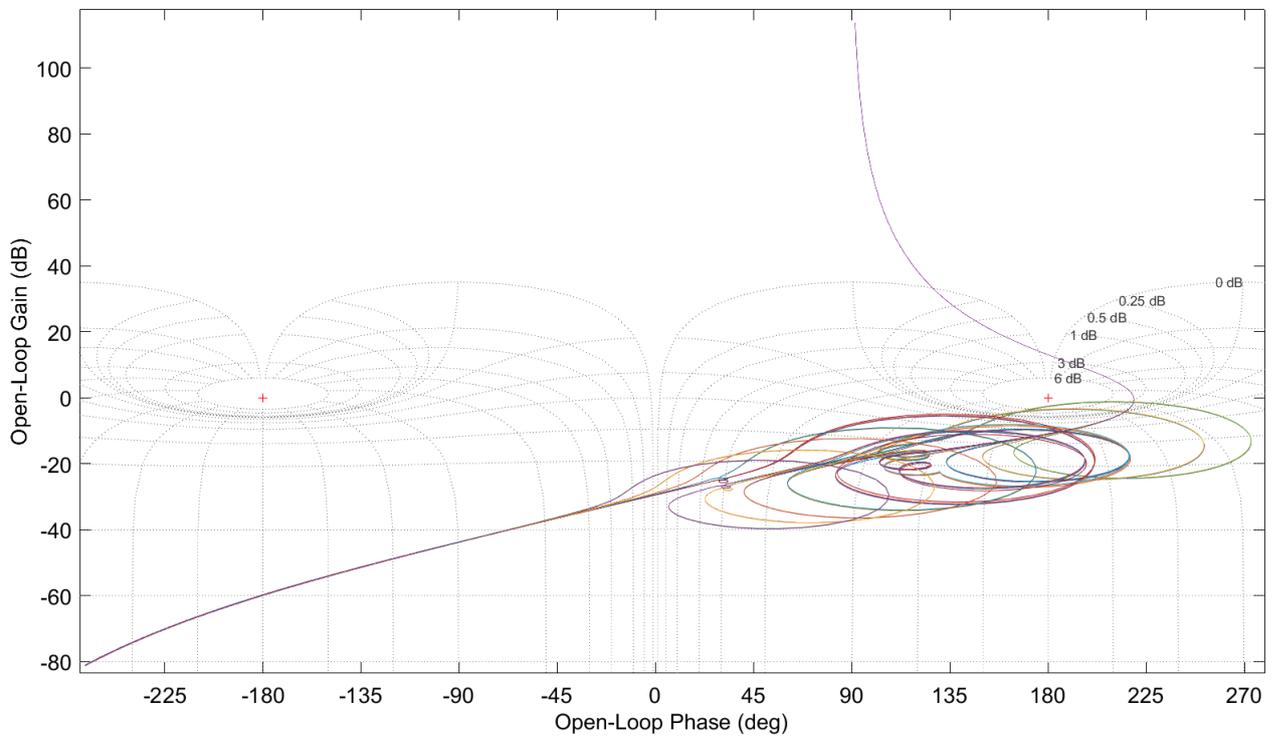


FIGURE 2.11 – Diagramme de Black de  $KG$  pour tout  $G \in \mathcal{B}$

## 2.7 Conclusion

La commande  $H_\infty$  par Loop-Shaping permet de concevoir un asservissement robuste à des incertitudes de modèle et à des perturbations extérieures, tout en considérant des contraintes de performance. Cette conception se réalise via la spécification de pondérations et la minimisation d'un critère  $H_\infty$  portant sur les quatre fonctions de sensibilité  $S$ ,  $GS$ ,  $KS$  et  $KGS$ .

La partie spécification de pondération est la partie la plus délicate et nécessite en pratique plusieurs itérations avant d'aboutir à un résultat acceptable. De plus, cette spécification n'est valide que pour une certaine incertitude de modèle spécifiée au préalable. Si ces spécifications d'incertitude de modélisation changent au cours d'un projet (par exemple lors d'une fabrication en série de longue durée), alors il est nécessaire de refaire le travail de conception de correcteur.

Ces problématiques amènent à entrevoir cette méthode de conception d'un angle de vue différent en considérant les paramètres d'un modèle comme indéterminés. C'est ce qu'on appelle dans cette thèse la *commande paramétrique*. L'objectif visé est de se donner de nouveaux outils afin de mieux comprendre le problème de commande et les systèmes que l'on étudie, tout ceci afin de gagner en efficacité lors de la conception de pondérations ou lors du calcul du correcteur.

Les outils du *calcul formel* que l'on introduit dans le chapitre suivant vont nous être très utiles dans cette tâche. Ils vont nous permettre de considérer des paramètres indéterminés dans notre problème de commande. Selon le type de systèmes étudiés, nous pourrions dire plus ou moins de choses sur le problème de commande mais, dans tous les cas, la formulation algébrique de ce problème apportera une plus-value intéressante par rapport à une résolution numérique classique.

Le chapitre suivant aborde la première difficulté du problème de commande  $H_\infty$  : la résolution de l'équation de Riccati  $\mathcal{R} = 0$ . Le Chapitre 4 utilise ensuite ces résultats pour l'étude du problème de commande  $H_\infty$ .

## Chapitre 3

# Approche symbolique des équations de Riccati algébriques

Le chapitre précédent a montré que les équations de Riccati algébriques (ERA) interviennent dans la résolution du Problème de Commande Robuste, tout comme dans la formule du problème de commande LQ ou LQG. Dans ce chapitre, nous étudions la résolution symbolique de telles équations. Par résolution symbolique, nous supposons que l'équation possède des *paramètres inconnus*, représentés sous forme symbolique, et que l'on cherche à réduire la résolution de cette équation à un système plus simple. Dans l'idéal, nous aimerions trouver une *paramétrisation* de l'ensemble des solutions, voire les représenter sous forme close.

L'avantage de résoudre symboliquement ces équations par rapport à des méthodes purement numériques (telles que, par exemple, celles décrites dans [BG96], [Lau79], ou [AL84] sur laquelle est basée la commande `gcare` de `MATLAB`) est que les solutions (ou le moyen de les représenter symboliquement) dépendent explicitement des paramètres. Ceci permet d'étudier, par exemple, comment ces solutions évoluent en fonction de ces paramètres. On parlera alors d'étude de *sensibilité par rapport aux paramètres*. Un autre avantage est de pouvoir certifier tout résultat, c'est-à-dire de délimiter complètement des zones singulières à certains problèmes dans l'espace des paramètres. Enfin, dans le cas où des formes closes sont calculables pour la solution d'un problème, un avantage est de pouvoir s'affranchir de tout algorithme d'optimisation pour résoudre le problème pour un jeu de paramètres donnés (valeurs numériques), puisque seule une évaluation de la forme close est alors nécessaire.

Dans la continuité des travaux de [FE93a, FE93b] ou [KYAH09], nous proposons une approche pour résoudre l'ERA utilisée pour la commande LQ/LQG/ $H_\infty$  de systèmes SISO sans transfert direct, que l'on note  $\mathcal{R} = 0$ . Après le rappel des définitions de quelques notions d'algèbre (Section 3.1), nous montrons comment réduire cette équation matricielle à un système d'équations polynomiales  $\mathcal{E}$  et ce par deux méthodes différentes (Section 3.3). Dans le cas où le système est considéré sous forme canonique, la première méthode est basée sur le système d'équation polynomiales découlant de l'ERA. Nous remarquons alors un schéma de substitution simplifiant grandement les calculs et ce, à n'importe quel ordre. La deuxième méthode est une généralisation de [KYAH09] dans le cas des systèmes à paramètres (en utilisant des techniques de la théorie des modules [Rot09]). Nous montrons alors l'équivalence des systèmes obtenus dans les deux méthodes. La paramétrisation des solutions de  $\mathcal{R} = 0$  suivant les solutions de  $\mathcal{E}$  est obtenue explicitement pour n'importe quel ordre  $n$ . En Section 3.4, nous détaillons les propriétés remarquables de ce système d'équations polynomiales. Ensuite, en Section 3.5, nous appliquons les outils présentés en Section 3.1 au système  $\mathcal{E}$  à paramètres fixés. Nous proposons alors une méthode afin de déterminer systématiquement

une paramétrisation des solutions de  $\mathcal{E}$ , puis d'identifier quelle solution de  $\mathcal{E}$  représente la solution définie positive de  $\mathcal{R} = 0$  en reprenant les travaux de [KYAH09]. En Section 3.6, nous étendons ces résultats au système  $\mathcal{E}$  à paramètres non fixés pour  $1 \leq n \leq 4$ , en se basant sur des calculs de cellules régulières et de variété discriminantes de  $\mathcal{E}$ . Ceci nous permet d'explicitier des formes closes pour les solutions de  $\mathcal{E}$ , et donc de  $\mathcal{R} = 0$ , quels que soient les paramètres, ainsi que d'identifier la solution définie positive parmi ces solutions. Enfin, en Section 3.7, nous détaillons quelques propriétés remarquables des systèmes comprenant un retard défini par une *approximation de Padé*.

Les résultats de ce chapitre sont publiés dans l'article de conférence [RBQQ16].

## Sommaire

---

<b>3.1</b>	<b>Quelques notions d'algèbre . . . . .</b>	<b>55</b>
3.1.1	Polynômes univariés et multivariés . . . . .	55
3.1.2	Extension algébrique et clôture algébrique . . . . .	55
3.1.3	Systèmes polynomiaux . . . . .	56
3.1.4	Classes d'équivalence . . . . .	58
3.1.5	Ordres monomiaux . . . . .	59
3.1.6	Division de polynômes à plusieurs variables . . . . .	60
3.1.7	Bases de Gröbner . . . . .	61
3.1.8	Matrices de multiplication . . . . .	63
3.1.9	Résultants . . . . .	64
3.1.10	Variétés discriminantes . . . . .	65
<b>3.2</b>	<b>Classe de systèmes et équation de Riccati algébrique <math>\mathcal{R} = 0</math> étudiés . . . . .</b>	<b>69</b>
<b>3.3</b>	<b>Réduction de <math>\mathcal{R} = 0</math> à un système d'équations polynomiales <math>\mathcal{B} = 0</math></b>	<b>70</b>
3.3.1	Paramétrisation des solutions de $\mathcal{R} = 0$ par un calcul des sous- espaces propres de la matrice Hamiltonienne associée . . . . .	70
3.3.2	Paramétrisation des solutions de $\mathcal{R} = 0$ par une étude directe du système d'équation polynomiales . . . . .	88
3.3.3	Équivalence des deux approches . . . . .	96
<b>3.4</b>	<b>Propriétés du système polynomial <math>\mathcal{E}</math> issu de <math>\mathcal{R} = 0</math> . . . . .</b>	<b>96</b>
<b>3.5</b>	<b>Étude de <math>\mathcal{E}</math> à paramètres fixés . . . . .</b>	<b>98</b>
3.5.1	Exemple d'introduction . . . . .	98
3.5.2	Éléments séparant $V(\langle \mathcal{B} \rangle)$ . . . . .	99
3.5.3	Paramétrisation de $V(\langle \mathcal{B} \rangle)$ . . . . .	103
3.5.4	Caractérisation de $X > 0$ solution de $\mathcal{R} = 0$ . . . . .	106
3.5.5	Isolation certifiée des zéros de $\mathcal{B}$ . . . . .	109
<b>3.6</b>	<b>Étude de <math>\mathcal{E}</math> à paramètres non fixés . . . . .</b>	<b>110</b>
3.6.1	Remarques préliminaires . . . . .	110
3.6.2	Paramétrisation des solutions de $\mathcal{E}$ pour $1 \leq n \leq 4$ . . . . .	111
3.6.3	Solutions de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$ . . . . .	113
3.6.4	Solution définie positive de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$ . . . . .	113
<b>3.7</b>	<b>Cas particulier des systèmes comportant une approximation de Padé . . . . .</b>	<b>122</b>
3.7.1	Rappels sur les approximations de Padé . . . . .	122
3.7.2	Approximations de Padé et factorisation spectrale . . . . .	123
<b>3.8</b>	<b>Conclusion . . . . .</b>	<b>124</b>

---

## 3.1 Quelques notions d'algèbre

### 3.1.1 Polynômes univariés et multivariés

Dans ce chapitre, nous étudions des polynômes définis sur un corps  $\mathbb{K}$ , tel que  $\mathbb{Q}$ ,  $\mathbb{R}$  ou  $\mathbb{C}$ . Le lecteur est certainement familier avec la notion de polynôme en une variable, aussi appelé *polynôme univarié*. En revanche, nous considérerons en général des polynômes de plusieurs variables. Nous parlerons alors de *polynôme multivarié*.

Afin d'être clair avec le formalisme que nous employons dans ce mémoire, nous énonçons quelques définitions générales.

**Définition 3.1.** Un *monôme* en les variables  $x_1, \dots, x_n$  est un produit de la forme  $x_1^{\alpha_1} \dots x_n^{\alpha_n}$  où les  $\alpha_i$  sont des entiers positifs, c'est-à-dire  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ . Nous noterons alors :

$$x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}.$$

**Définition 3.2.** Un *polynôme*  $f$  en les variables  $x_1, \dots, x_n$  et à coefficients dans un corps  $\mathbb{K}$  est une combinaison  $\mathbb{K}$ -linéaire finie de monômes :

$$f = \sum_{0 \leq |\alpha| \leq r} a_\alpha x^\alpha, \quad a_\alpha \in \mathbb{K}, \quad |\alpha| = \alpha_1 + \dots + \alpha_n, \quad r \in \mathbb{N}.$$

L'ensemble des polynômes en  $x_1, \dots, x_n$  et à coefficients dans  $\mathbb{K}$  est noté par  $\mathbb{K}[x_1, \dots, x_n]$ .

**Remarque 3.1.**  $(\mathbb{K}[x_1, \dots, x_n], +, \times)$  est un *anneau commutatif* (par abus de notation, nous parlerons simplement de l'anneau  $\mathbb{K}[x_1, \dots, x_n]$ ). Pour des rappels sur les anneaux, voir, par exemple, [FED09].

Lorsque  $\mathbb{K} = \mathbb{C}$ , un résultat classique, mais non moins utile, est le suivant.

**Théorème 3.1.** Tout polynôme non constant  $f \in \mathbb{C}[x]$  admet une racine dans  $\mathbb{C}$ .

Si l'on veut étendre ce résultat à d'autres corps que  $\mathbb{C}$ , il est nécessaire de définir la notion de *clôture algébrique*.

### 3.1.2 Extension algébrique et clôture algébrique

Soient  $\mathbb{K}$  et  $\mathbb{L}$  deux corps tels que  $\mathbb{K} \subseteq \mathbb{L}$ . On dira alors que  $\mathbb{L}$  est une *extension algébrique* de  $\mathbb{K}$  si tout élément  $l$  de  $\mathbb{L}$  satisfait une équation polynomiale univariée à coefficients dans  $\mathbb{K}$ , c'est-à-dire qu'il existe  $p \in \mathbb{K}[x]$  tel que  $p(l) = 0$ . De plus, cette extension algébrique sera dite *algébriquement close* si toute solution d'équation polynomiale univariée à coefficient dans  $\mathbb{K}$  appartient à  $\mathbb{L}$ , c'est-à-dire  $p(x) = 0$  implique que  $x \in \mathbb{L}$  pour tout  $p \in \mathbb{K}[x]$ . On peut montrer que tout corps  $\mathbb{K}$  admet une clôture algébrique notée  $\overline{\mathbb{K}}$  [Lan02].

Par exemple,  $\mathbb{Q}(\sqrt[3]{2}) := \{a + b\sqrt[3]{2} \mid (a, b) \in \mathbb{Q}^2\}$  est une extension algébrique du corps  $\mathbb{Q}$  qui n'est pas algébriquement close car  $j\sqrt[3]{2}$  et  $-j\sqrt[3]{2}$ , racines de  $x^3 - 2$ , n'appartiennent pas à  $\mathbb{Q}(\sqrt[3]{2})$ .  $\mathbb{C}$  est une clôture algébrique de  $\mathbb{R}$  puisque tout élément  $z = a + ib$  de  $\mathbb{C}$ ,  $(a, b) \in \mathbb{R}^2$ , est racine du polynôme  $x^2 - 2ax + a^2 + b^2$  à coefficients réels et toute racine d'un polynôme à coefficients réels est dans  $\mathbb{C}$ .

### 3.1.3 Systèmes polynomiaux

Dans la suite de ce mémoire, nous considérons des systèmes d'équations polynomiales. Un système d'équation polynomiales est défini par des polynômes  $f_1, \dots, f_s$  de  $\mathbb{K}[x_1, \dots, x_n]$  tels que :

$$\begin{cases} f_1(x_1, \dots, x_n) = 0, \\ \vdots \\ f_s(x_1, \dots, x_n) = 0. \end{cases} \quad (3.1)$$

**Exemple 3.1.** Soit  $\mathbb{K} = \mathbb{Q}$ . Un exemple de système d'équations polynomiales de  $\mathbb{Q}[x_1, \dots, x_n]$  est :

$$\begin{cases} x_1^2 + 2x_1x_2 + x_2^2 = 0, \\ x_1^2x_2^2 - 1 = 0. \end{cases}$$

Soit  $\mathbb{K} = \mathbb{Q}[a]$ . Un exemple de système d'équations polynomiales de  $\mathbb{Q}[a][x_1, \dots, x_n]$  est :

$$\begin{cases} ax_1^2 + 2x^2 = 0, \\ x_2^2x_1 + 2x_1^2 - a^2 = 0. \end{cases}$$

A ces systèmes de polynômes, nous associons un *idéal*. Rappelons tout d'abord cette notion.

**Définition 3.3.** Soit  $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ .  $I$  est un *idéal* si :

- $(I, +)$  est un sous-groupe de  $(\mathbb{K}[x_1, \dots, x_n], +)$ ,
- Si  $f \in I$  et  $p \in \mathbb{K}[x_1, \dots, x_n]$ , alors  $pf \in I$ .

Dans le cas de systèmes polynomiaux, nous associons l'idéal suivant.

**Définition 3.4.** Soient  $(f_1, \dots, f_s) \in \mathbb{K}[x_1, \dots, x_n]^s$ . Alors, nous notons :

$$\langle f_1, \dots, f_s \rangle := \left\{ \sum_{i=1}^s p_i f_i \mid (p_1, \dots, p_s) \in \mathbb{K}[x_1, \dots, x_n]^s \right\}.$$

Alors  $\langle f_1, \dots, f_s \rangle$  est un idéal, appelé *idéal de l'anneau*  $\mathbb{K}[x_1, \dots, x_n]$  *engendré/généré par*  $f_1, \dots, f_s$ .

**Remarque 3.2.** L'idéal  $\langle f_1, \dots, f_s \rangle$  est même le plus petit idéal de  $\mathbb{K}[x_1, \dots, x_n]$  contenant  $\{f_1, \dots, f_s\}$ .

Lorsque  $\mathbb{K} = \mathbb{C}$ , nous pouvons représenter géométriquement les solutions de (3.1). Nous parlerons alors de *variété algébrique complexe*. Nous étendrons cette notion sur un corps  $\mathbb{K}$  quelconque, bien que les solutions soient généralement définies dans la clôture algébrique  $\overline{\mathbb{K}}$ . Les variétés algébriques en lien avec nos problèmes de commande sont centrales dans cette thèse.

**Définition 3.5.** Soit  $\mathcal{A}$  une  $\mathbb{K}$ -algèbre. Soient  $f_1, \dots, f_s$  des polynômes de  $\mathbb{K}[x_1, \dots, x_n]$  et  $\langle f_1, \dots, f_s \rangle$  l'idéal associé. Alors, nous définissons :

$$V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle) := \{(\nu_1, \dots, \nu_n) \in \mathcal{A}^n \mid \forall p \in \langle f_1, \dots, f_s \rangle, p(\nu_1, \dots, \nu_n) = 0\}.$$

$V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle)$  est appelé *ensemble algébrique défini par*  $f_1, \dots, f_s$  *et*  $\mathcal{A}$ , ou encore *variété algébrique affine (formée des  $\mathcal{A}$ -points) de*  $\langle f_1, \dots, f_s \rangle$ . Dans la suite, on prendra  $\mathcal{A} = \overline{\mathbb{K}}$  et on notera  $V(\langle f_1, \dots, f_s \rangle) := V_{\overline{\mathbb{K}}}(\langle f_1, \dots, f_s \rangle)$ , qui sera simplement appelée *variété*.

**Remarque 3.3.** Dans le cas où  $\mathbb{K} = \mathbb{C}$ , la variété  $V_{\mathbb{C}}(\langle f_1, \dots, f_s \rangle)$  est un ensemble algébrique de  $\mathbb{C}^n$  définie par les zéros communs aux  $f_i$ .

**Remarque 3.4.** Reprenons les notations de la Définition 3.5. Si un polynôme  $f \in \mathbb{K}[x_1, \dots, x_n]$  s'écrit

$$f = r + \sum_{i=1}^s q_i f_i, \quad r, q_i \in \mathbb{K}[x_1, \dots, x_n],$$

alors  $f = r$  sur la variété  $V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle)$ .

Nous définissons aussi un idéal associé à une variété algébrique comme suit.

**Définition 3.6.** Soit  $V$  une variété algébrique. Alors, nous notons  $I(V)$  l'ensemble des polynômes s'annulant sur la variété  $V$ , c'est-à-dire :

$$I(V) := \{f \in \mathbb{K}[x_1, \dots, x_n] \mid \forall (\nu_1, \dots, \nu_n) \in V, f(\nu_1, \dots, \nu_n) = 0\}.$$

$I(V)$  est un idéal, appelé *idéal de  $V$* .

Nous disposons de plusieurs objets pour caractériser un système polynomial. Il existe des correspondances entre ces objets. Notamment, comme tout élément  $p$  de  $\langle f_1, \dots, f_s \rangle$  s'écrit sous la forme de  $p_1 f_1 + \dots + p_s f_s$ , où  $(p_1, \dots, p_s) \in \mathbb{K}[x_1, \dots, x_n]^s$ , nous avons :

$$E := \{\nu = (\nu_1, \dots, \nu_n) \in \mathcal{A}^n \mid f_i(\nu_1, \dots, \nu_n) = 0, i = 1, \dots, s\} \subseteq V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle).$$

Réciproquement, si  $\nu \in V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle)$ , en utilisant que  $f_i \in \langle f_1, \dots, f_s \rangle$  pour  $i = 1, \dots, s$ , nous obtenons que  $f_i(\nu) = 0$  pour  $i = 1, \dots, s$ , c'est-à-dire que  $\nu \in E$  et donc que  $V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle) \subseteq E$ , ce qui prouve que  $E = V_{\mathcal{A}}(\langle f_1, \dots, f_s \rangle)$ .

**Proposition 3.1.** Si  $f_1, \dots, f_s$  et  $f'_1, \dots, f'_t$  sont deux familles génératrices d'un idéal  $I$ , c'est-à-dire que  $I = \langle f_1, \dots, f_s \rangle = \langle f'_1, \dots, f'_t \rangle$ , alors nous avons :  $V(\langle f_1, \dots, f_s \rangle) = V(\langle f'_1, \dots, f'_t \rangle)$ , simplement noté par  $V(I)$ .

L'idéal contient une information supplémentaire que la variété algébrique n'encode pas : la multiplicité. Par exemple, considérons les système suivants :

- $x^2 = 0$  : l'idéal associé est  $I_1 = \langle x^2 \rangle$ .
- $x = 0$  : l'idéal associé est  $I_2 = \langle x \rangle$ .

Nous avons  $I_2 \not\subseteq I_1$ . Par contre, ces deux idéaux sont associés à la même variété algébrique  $V(I_1) = V(I_2) = \{0\}$ . En d'autres termes, la *multiplicité* des solutions d'un système d'équation polynomiales n'est pas prise en compte dans la variété algébrique associée à celui-ci. Autrement dit, si un polynôme multivarié s'annule sur tous les points d'une variété  $V(I)$ , ceci n'implique pas nécessairement que  $f \in I$ .

**Théorème 3.2.** (*Hilbert Nullstellensatz*) Soient  $\mathcal{A}$  une  $\mathbb{K}$ -algèbre et  $f, f_1, \dots, f_s \in \mathbb{K}[x_1, \dots, x_n]^{s+1}$ , où  $\mathbb{K}$  est un corps algébriquement clos tel que  $\mathbb{C}$ . Alors, nous avons :

$$f \in I(V_{\mathbb{K}}(\langle f_1, \dots, f_s \rangle)) \Rightarrow \exists k \in \mathbb{N}, f^k \in \langle f_1, \dots, f_s \rangle.$$

Nous définissons alors la notion d'idéal radical.

**Définition 3.7.** Si  $I$  est un idéal de  $\mathbb{K}[x_1, \dots, x_n]$ , alors nous définissons le *radical* de  $I$ , noté  $\sqrt{I}$ , par l'ensemble :

$$\sqrt{I} := \{f \in \mathbb{K}[x_1, \dots, x_n] \mid f^m \in I, m \geq 1\}.$$

Un idéal  $I$  est dit *radical* si  $\sqrt{I} = I$ .

Afin de déterminer certaines propriétés de la variété  $V(\langle f_1, \dots, f_s \rangle)$  (par exemple sa dimension), on est amené à changer de générateurs de l'idéal  $\langle f_1, \dots, f_s \rangle$  en considérant, par exemple, une *base de Gröbner*. Nous évoquerons cette notion dans les sections qui suivent.

Notons aussi que, disposant du système d'équations (3.1), nous pouvons "simplifier" tout polynôme dans  $\mathbb{K}[x_1, \dots, x_n]$  par  $(f_1, \dots, f_s)$  en utilisant le concept de *classe d'équivalence*.

### 3.1.4 Classes d'équivalence

**Définition 3.8.** Soit  $r \in \mathbb{K}[x_1, \dots, x_n]$ . On appelle *classe d'équivalence* de  $r$  modulo  $\langle f_1, \dots, f_s \rangle$  l'ensemble de polynômes défini par :

$$\bar{r} := \{\alpha \in \mathbb{K}[x_1, \dots, x_n] \mid r - \alpha \in \langle f_1, \dots, f_s \rangle\}.$$

$r$  est dit *représentant* de la classe d'équivalence  $\bar{r}$ .

L'ensemble des classes d'équivalence des polynômes de  $\mathbb{K}[x_1, \dots, x_n]$  modulo  $\langle f_1, \dots, f_s \rangle$  est un anneau.

**Théorème 3.3.** [CLO15, Th. 6, Chap. 5, §2] L'ensemble des classes d'équivalences des polynômes de  $\mathbb{K}[x_1, \dots, x_n]$  modulo  $\langle f_1, \dots, f_s \rangle$  est un anneau pour les opérations suivantes

$$\forall (r_1, r_2) \in \mathbb{K}[x_1, \dots, x_n]^2, \quad \overline{r_1 + r_2} = \overline{r_1} + \overline{r_2}, \quad \overline{r_1 r_2} = \overline{r_1} \overline{r_2},$$

appelé *anneau quotient* de  $\mathbb{K}[x_1, \dots, x_n]$  par  $\langle f_1, \dots, f_s \rangle$  et noté par  $\mathcal{A} := \mathbb{K}[x_1, \dots, x_n] / \langle f_1, \dots, f_s \rangle$ .

**Exemple 3.2.** Considérons l'idéal  $\langle f_1 \rangle$  où  $f_1 = x_1^2 - 1$ . Le polynôme  $g = x_1^3$  peut se réduire modulo  $\langle f_1 \rangle$ . En effet, nous avons

$$f_1 = x_1^2 - 1 = 0,$$

ce qui implique :

$$g = x_1^3 = x_1.$$

Ainsi,  $x_1^3$  et  $x_1$  appartiennent à la même classe d'équivalence modulo  $f_1$ , c'est-à-dire définissent le même élément dans le quotient  $\mathbb{K}[x_1] / \langle f_1 \rangle$ .

Notons que l'anneau quotient  $\mathcal{A}$  vérifie les propriétés suivantes.

**Proposition 3.2.** [FED09]  $\mathcal{A}$  est une  $\mathbb{K}$ -algèbre. En particulier,  $\mathcal{A}$  est un  $\mathbb{K}$ -espace-vectoriel. De plus,  $\mathcal{A}$  est engendré comme  $\mathbb{K}$ -espace-vectoriel par l'ensemble des classes d'équivalence des monômes de  $\mathbb{K}[x_1, \dots, x_n]$ .

Nous accorderons une importance particulière aux anneaux quotients zéro-dimensionnels définis comme suit.

**Définition 3.9.**  $\mathcal{A}$  est dit *zéro-dimensionnel* si la dimension de  $\mathcal{A}$  en tant que  $\mathbb{K}$  espace vectoriel est finie. Par abus de langage, on dira que  $\langle f_1, \dots, f_s \rangle$  est zéro-dimensionnel.

Il existe des algorithmes permettant de calculer un représentant “simple” d’une classe d’équivalence, comme par exemple calculer une *forme normale* [CLO15, Prop. 1, Chap. 2, §6]. Calculer une forme normale nécessite, par exemple, la connaissance d’une *base de Gröbner* de l’idéal  $\langle f_1, \dots, f_s \rangle$  engendré par ce système pour un *ordre monomial* donné. Une base de Gröbner de  $\langle f_1, \dots, f_s \rangle$  est une famille génératrice de l’idéal  $\langle f_1, \dots, f_s \rangle$  possédant certaines propriétés supplémentaires que nous détaillerons dans les sections suivantes. Disposant d’une base de Gröbner de  $\langle f_1, \dots, f_s \rangle$ , nous pouvons alors effectuer des opérations effectives dans l’anneau quotient  $\mathcal{A} := \mathbb{K}[x_1, \dots, x_n]/\langle f_1, \dots, f_s \rangle$ , en particulier vérifier si  $r \in \langle f_1, \dots, f_s \rangle$  en testant que la forme normale  $\bar{r}$  de  $r$  vérifie  $\bar{r} = 0$  (la définition de la forme normale est explicitée en Section 3.1.7).

### 3.1.5 Ordres monomiaux

Nous avons vu qu’il y avait une correspondance entre idéaux et systèmes polynomiaux. Il peut être intéressant de trouver une autre base génératrice de l’idéal  $I$  pour “résoudre” certains problèmes comme reconnaître si un polynôme appartient à un idéal  $I$ , ou rechercher les solutions du système polynomial défini par un idéal zéro-dimensionnel. Les bases de Gröbner permettent de répondre à ces questions. Avant de définir celles-ci, nous devons parler de division de polynômes multivariés et d’un algorithme permettant de diviser un polynôme par un ensemble de polynômes. Dans le cas univarié, celui-ci s’appelle *division euclidienne*. Cet algorithme se base sur l’ordonnancement des termes d’un polynôme suivant leur degré. Si l’on cherche à étendre cet algorithme dans le cas multivarié, nous devons définir la notion d’ordre pour des polynômes à plusieurs variables. C’est ce qui est appelé *ordre monomial*.

Nous donnons ici la définition d’un ordre monomial ainsi que quelques exemples. Pour plus de détails, le lecteur est invité à se référer à [CLO15, Chap.2, §2].

**Définition 3.10.** Un *ordre monomial*  $\succ$  sur  $\mathbb{K}[x_1, \dots, x_n]$  est une relation  $\succ$  sur  $\mathbb{N}^n$  (ou, de manière équivalente, une relation sur l’ensemble des monômes  $x^\alpha$ ,  $\alpha \in \mathbb{N}^n$ ), satisfaisant :

- $\succ$  est un *ordre total* sur  $\mathbb{N}^n$ , c’est à dire une relation d’ordre pour laquelle deux éléments de  $\mathbb{N}^n$  sont toujours comparables,
- si  $\alpha \succ \beta$  et  $\gamma \in \mathbb{N}^n$ , alors  $\alpha + \gamma \succ \beta + \gamma$ ,
- chaque sous-ensemble non vide de  $\mathbb{N}^n$  a un plus petit élément pour l’ordre  $\succ$ .

Il existe plusieurs exemples d’ordres monomiaux. Nous en donnons quelques uns ici.

**Définition 3.11.** Soient  $\alpha = (\alpha_1, \dots, \alpha_n)$  et  $\beta = (\beta_1, \dots, \beta_n)$  dans  $\mathbb{N}^n$ . Nous disons que  $\alpha \succ_{lex} \beta$  si la première composante non nulle de  $\alpha - \beta$  est positive. Nous écrivons alors  $x^\alpha \succ_{lex} x^\beta$  si  $\alpha \succ_{lex} \beta$ . Cet ordre est appelé *ordre lexicographique*.

**Définition 3.12.** Soient  $\alpha = (\alpha_1, \dots, \alpha_n)$  et  $\beta = (\beta_1, \dots, \beta_n)$  dans  $\mathbb{N}^n$ . Rappelons que  $|\alpha| = \alpha_1 + \dots + \alpha_n$ . Nous disons que  $\alpha \succ_{DRL} \beta$  si  $|\alpha| > |\beta|$  ou si  $|\alpha| = |\beta|$  et la dernière composante non-nulle de  $\alpha - \beta$  est négative. Nous écrivons alors  $x^\alpha \succ_{DRL} x^\beta$  si  $\alpha \succ_{DRL} \beta$ . Cet ordre est appelé *ordre du Degré Lexicographique Inverse*, abrégé *DRL* (*Degree-Reverse-Lexicographical*).

**Exemple 3.3.** Considérons les monômes  $x_1^3 x_3^2$  et  $x_1^2 x_2^2 x_3$ . Alors, nous avons  $\alpha = (3, 0, 2)$  et  $\beta = (2, 2, 1)$  et  $\alpha - \beta = (1, -2, 1)$ . Ainsi,  $x_1^3 x_3^2 \succ_{lex} x_1^2 x_2^2 x_3$ . Par ailleurs,  $\sum_{i=1}^3 \alpha_i = \sum_{i=1}^3 \beta_i = 5$ , d’où  $x_1^3 x_3^2 \succ_{DRL} x_1^2 x_2^2 x_3$ .

Disposant d'un ordre monomial  $\succ$ , nous pouvons alors ordonner les polynômes de  $\mathbb{K}[x_1, \dots, x_n]$ . Nous définissons les notions suivantes.

**Définition 3.13.** Soit  $f := \sum_{\alpha} a_{\alpha} x^{\alpha}$  un polynôme non nul de  $\mathbb{K}[x_1, \dots, x_n]$  et  $\succ$  un ordre monomial.

- Le *degré* de  $f$  est :

$$\deg(f) = \max \{ \alpha \in \mathbb{N}^n \mid a_{\alpha} \neq 0 \}.$$

- Le *degré total* de  $f$  est  $\max \{ |\alpha| \mid a_{\alpha} \neq 0 \}$ .
- Le *coefficient de tête* (*leading coefficient* en anglais) de  $f$  est :

$$\text{LC}(f) = a_{\deg(f)} \in \mathbb{K}.$$

- Le *monôme de tête* (*leading monomial* en anglais) de  $f$  est :

$$\text{LM}(f) = x^{\deg(f)}.$$

- Le *terme de tête* (*leading term* en anglais) de  $f$  est :

$$\text{LT}(f) = \text{LC}(f) \text{LM}(f).$$

**Remarque 3.5.** Lorsque le contexte est clair, nous ne précisons par l'ordre monomial utilisé.

**Remarque 3.6.** Si  $n = 1$ , alors les polynômes de  $\mathbb{K}[x_1]$  sont ordonnés par leur degré.

### 3.1.6 Division de polynômes à plusieurs variables

Disposant d'un ordre monomial  $\succ$ , nous pouvons définir un *algorithme de division* dans le cas multivarié. Étant donné  $(f, f_1, \dots, f_s) \in \mathbb{K}[x_1, \dots, x_n]^{s+1}$ , l'algorithme calcule  $(r, q_1, \dots, q_s) \in \mathbb{K}[x_1, \dots, x_n]^{s+1}$  tel que :

$$f = r + \sum_{i=1}^s q_i f_i.$$

Un théorème en lien avec un tel algorithme est le suivant.

**Théorème 3.4.** Soient  $\succ$  un ordre monomial sur  $\mathbb{N}^n$  et  $F = (f_1, \dots, f_s)$  un  $s$ -uplet de polynômes ordonné de  $\mathbb{K}[x_1, \dots, x_n]$ . Alors, tout polynôme  $f \in \mathbb{K}[x_1, \dots, x_n]$  peut s'écrire

$$f = q_1 f_1 + \dots + q_s f_s + r, \tag{3.2}$$

où  $q_i, r \in \mathbb{K}[x_1, \dots, x_n]$  et soit  $r = 0$ , soit  $r$  est une combinaison  $\mathbb{K}$ -linéaire de monômes non divisibles par les termes de tête des  $f_i$ . Alors,  $r$  est appelé *reste* de la division de  $f$  par  $F$ .

Ce théorème se prouve simplement en explicitant un algorithme permettant de calculer les  $q_i$  et  $r$ . L'algorithme est le suivant.

**Algorithme 3.1.** (division de  $f$  par un  $s$ -uplet ordonné  $(f_1, \dots, f_s)$ )

Entrées :  $(f_1, \dots, f_s) \in \mathbb{K}[x_1, \dots, x_n]^s$  un  $s$ -uplet ordonné,  $f \in \mathbb{K}[x_1, \dots, x_n]$   
 et  $\succ$  un ordre monomial

Sorties :  $q_1, \dots, q_s, r \in \mathbb{K}[x_1, \dots, x_n]$  tels que (3.2) et  $r$  est non divisible par  
 les termes de tête des  $f_i$ ,  $i = 1, \dots, s$

$q_1 := 0; \dots; q_s := 0; r := 0;$   
 $p := f;$   
 tant que  $p \neq 0$  faire  
    $i := 1$   
   division\_effectuée := *faux*  
   tant que  $i \leq s$  et division\_effectuée = *faux* faire  
     si  $\text{LT}(f_i)$  divise  $\text{LT}(p)$  alors  
        $q_i := q_i + \text{LT}(p)/\text{LT}(f_i)$   
        $p := p - (\text{LT}(p)/\text{LT}(f_i)) f_i$   
       division\_effectuée := *vrai*  
     sinon  
        $i := i + 1$   
   si division\_effectuée = *faux* alors  
      $r := r + \text{LT}(p)$   
      $p := p - \text{LT}(p)$   
 retourner  $q_1, \dots, q_s, r$

Notons que, lorsque  $\text{LT}(f_i)$  divise  $\text{LT}(p)$ , alors l'algorithme annule des termes de tête. Notons aussi que, dans le cas  $n = 1$ , nous retrouvons l'algorithme de division euclidienne.

Il est important de noter que suivant l'ordre des  $f_i$  dans lequel l'algorithme de division est réalisé, celui-ci peut retourner des résultats différents.

**Exemple 3.4.** [CLO15, Chap. 2] Soient  $f = x^2 y + x y^2 + y^2$ ,  $f_1 = y^2 - 1$  et  $f_2 = x y - 1$ . La division de  $f$  par  $(f_1, f_2)$  suivant l'ordre  $x \succ_{lex} y$  retourne le reste  $2x + 1$ . En changeant l'ordre des polynômes diviseurs, c'est-à-dire en divisant  $f$  par  $(f_2, f_1)$  suivant l'ordre  $x \succ_{lex} y$ , nous obtenons le reste  $x$ , différent du reste précédent.

Si  $\{f_1, \dots, f_s\}$  possède certaines propriétés, alors le résultat de l'algorithme est unique. Les bases de Gröbner, définies à la section suivante, vérifient une telle propriété.

### 3.1.7 Bases de Gröbner

Une base de Gröbner d'un idéal  $I$  est définie comme suit.

**Définition 3.14.** Soit un ordre monomial  $\succ$ . Un sous ensemble fini  $G = \{g_1, \dots, g_t\} \neq \{0\}$  d'un idéal  $I \subseteq \mathbb{K}[x_1, \dots, x_n]$  est appelée *base de Gröbner* si :

$$\langle \text{LT}(g_1), \dots, \text{LT}(g_t) \rangle = \langle \text{LT}(I) \rangle.$$

Cette définition étant difficilement applicable en pratique, nous retenons généralement les propriétés des bases de Gröbner.

**Théorème 3.5.** Soient  $I \subseteq \mathbb{K}[x_1, \dots, x_n]$  un idéal,  $G = \{g_1, \dots, g_t\}$  une base de Gröbner de  $I$  pour l'ordre  $\succ$  et  $f \in \mathbb{K}[x_1, \dots, x_n]$ . Alors, il existe un unique  $r \in \mathbb{K}[x_1, \dots, x_n]$  avec les propriétés suivantes :

1. Il existe  $g \in I$  tel que  $f = g + r = \sum_{i=1}^t g_i q_i + r$ , pour  $g_1, \dots, g_t \in \mathbb{K}[x_1, \dots, x_n]$ .
2. Aucun terme de  $r$  n'est divisible par  $LT(g_i)$  pour  $1 \leq i \leq t$ .

En particulier,  $r$  est l'unique reste de la division de  $f$  par  $G$ .

Soient  $I \subseteq \mathbb{K}[x_1, \dots, x_n]$  un idéal et  $G = \{g_1, \dots, g_t\}$  une base de Gröbner de  $I$ . L'algorithme de division (voir Algorithme 3.1) fournit alors un reste unique lorsque l'on divise le polynôme  $f$  par  $G$ . Ce reste est alors appelé *forme normale* et est noté  $\overline{f}^G$ . Cet algorithme permet notamment de calculer un représentant "simple" (réduit) d'une classe d'équivalence.

**Exemple 3.5.** Reprenons l'Exemple 3.4. Notons  $I := \langle f_1, f_2 \rangle$  avec  $f_1 := y^2 - 1$  et  $f_2 := xy - 1$ . Une base de Gröbner de  $I$  pour l'ordre lexicographique  $x \succ_{lex} y$  est  $G_1 := \{y^2 - 1, x - y\}$ , tandis qu'une pour l'ordre  $x \succ_{DRL} y$  est  $G_2 := \{x - y, y^2 - 1\}$ . Le reste de la division de  $f$  par  $G_1$  pour l'ordre  $x \succ_{lex} y$  ou par  $G_2$  pour l'ordre  $x \succ_{DRL} y$  est identique et vaut  $2y + 1$ .

Tout d'abord, nous aimerions savoir reconnaître quand un ensemble de polynômes forme une base de Gröbner, et ensuite, dans le cas négatif, nous voudrions calculer la base de Gröbner associée à cet ensemble. Pour cela, nous avons besoin d'introduire la définition suivante.

**Définition 3.15.** [CLO15, Chap. 2, §6] Soient  $f$  et  $g$  dans  $\mathbb{K}[x_1, \dots, x_n]$ . Le *S-polynôme* de  $f$  et  $g$  est alors la combinaison suivante

$$S(f, g) = \frac{x^\gamma}{LT(f)} f - \frac{x^\gamma}{LT(g)} g,$$

où  $x^\gamma$  désigne le *plus petit commun multiple* de  $LM(f)$  et  $LM(g)$ , c'est-à-dire  $\gamma = (\gamma_1, \dots, \gamma_n)$ , avec  $\gamma_i = \max\{\alpha_i, \beta_i\}$  avec  $\alpha = \deg(f)$  et  $\beta = \deg(g)$ .

Le critère suivant permet alors d'identifier une base de Gröbner.

**Théorème 3.6.** (*Critère de Buchberger*) [CLO15, Th. 6, §6, Chap. 2] Soit  $I$  un idéal polynomial. Alors, une famille génératrice  $G = \{g_1, \dots, g_t\}$  de  $I$  est une base de Gröbner de  $I$  si et seulement si pour toute paire  $(i, j)$  avec  $i \neq j$ , le reste de la division de  $S(g_i, g_j)$  par  $G$  est zéro, c'est-à-dire  $\forall i \neq j, \overline{S(g_i, g_j)}^G = 0$ .

L'algorithme suivant permet de calculer une base de Gröbner d'un idéal  $I = \langle f_1, \dots, f_s \rangle \neq \{0\}$  en un nombre fini d'étapes.

**Algorithme 3.2.** (*Algorithme de Buchberger*) [CLO15, Chap. 2, §7]

Entrées :  $F := \{f_1, \dots, f_s\}$  tel que  $I = \langle f_1, \dots, f_s \rangle$   
 Sorties : Une base de Gröbner  $G = (g_1, \dots, g_s)$  de  $I$

$G := F$ ;  
 répéter  
      $G' := G$   
     pour toute paire  $\{p, q\}, p \neq q$  dans  $G'$  faire  
          $r := \overline{S(p, q)}^{G'}$   
         si  $r \neq 0$  alors  $G := G \cup \{r\}$   
 jusqu'à  $G = G'$   
 retourner  $G$

Pour des raffinements sur cet algorithme, voir, par exemple, [CLO15, Chap. 2].

Afin de déterminer si deux ensembles de polynômes génèrent le même idéal, il est intéressant de disposer d'une base de Gröbner réduite. Celle-ci est définie comme suit.

**Définition 3.16.** Une *base de Gröbner réduite* d'un idéal polynomial  $I$  est une base de Gröbner  $G$  de  $I$  telle que :

- $\text{LC}(p) = 1$  pour tout  $p \in G$ ,
- pour tout  $p \in G$  aucun monôme de  $p$  n'est divisible par un des monômes de  $G \setminus \{0\}$ .

Une base de Gröbner réduite possède la propriété suivante.

**Théorème 3.7.** Soit  $I \neq \{0\}$  un idéal polynomial. Alors, disposant d'un ordre monomial  $\succ$ ,  $I$  admet une unique base de Gröbner réduite.

Le Théorème 3.7 permet de montrer que deux ensembles de polynômes génèrent le même idéal.

### 3.1.8 Matrices de multiplication

Soit  $f := \{f_1, \dots, f_s\}$  un système de polynômes de  $R := \mathbb{K}[x_1, \dots, x_n]$ . Soit  $I$  l'idéal généré par  $f$ , que l'on suppose zéro-dimensionnel. Notons par  $\mathcal{A} := R/I$  l'anneau quotient de  $R$  par  $I$ . Nous définissons la notion de *matrice de multiplication* de la manière suivante.

**Définition 3.17.** Soit  $r \in R$ . Le  $\mathbb{K}$ -endomorphisme de multiplication de  $\mathcal{A}$  associé à  $r \in R$  est défini par :

$$\begin{aligned} M_r : \mathcal{A} &\longrightarrow \mathcal{A} \\ \bar{a} &\longmapsto \overline{ra}. \end{aligned}$$

Autrement dit, l'image de  $\bar{a}$  par cet endomorphisme est la réduction de  $ra$  modulo  $I$ . Alors, on peut exprimer  $\bar{a}$  et  $\overline{ra}$  dans une  $\mathbb{K}$ -base de  $\mathcal{A}$ .  $m_r$  est alors défini par une matrice, permettant de passer du vecteur coordonnée de  $\bar{a}$  à celui de  $\overline{ra}$ ; nous parlons alors de *matrice de multiplication*.

**Exemple 3.6.** Notons  $\mathbb{K} := \mathbb{Q}(d_0, d_2)$  le corps des fractions rationnelles en les paramètres  $d_0, d_2$  et  $R := \mathbb{K}[b_0, b_1]$  l'ensemble des polynômes en les variables  $b_0$  et  $b_1$  et à coefficients dans  $\mathbb{K}$ . Soit  $\mathcal{B}$  un ensemble de polynômes de  $R$  défini par :

$$\mathcal{B} := \{b_0^2 - d_0, b_1^2 - 2b_0 - d_2\}.$$

Cet ensemble de polynômes sera étudié par la suite. Notons par  $\mathcal{A} := R/\langle \mathcal{B} \rangle$  l'anneau quotient de  $R$  par  $\langle \mathcal{B} \rangle$ . Une  $\mathbb{K}$ -base de  $\mathcal{A}$  est  $\{1, b_1, b_0, b_1 b_0\}$ . La matrice de multiplication par  $b_0$ , notée par  $M_{b_0}$ , a la forme suivante :

$$M_{b_0} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ d_0 & 0 & 0 & 0 \\ 0 & d_0 & 0 & 0 \end{pmatrix}.$$

Celle-ci vérifie bien :

$$\begin{pmatrix} b_0 \\ b_0 b_1 \\ b_0^2 \\ b_1 b_0^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ d_0 & 0 & 0 & 0 \\ 0 & d_0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ b_1 \\ b_0 \\ b_1 b_0 \end{pmatrix}.$$

Soit  $\chi = (\chi_1, \dots, \chi_n) \in \overline{\mathbb{K}}^n$ . Soit  $r(x_1, \dots, x_n) \in R$ . Nous notons  $r(\chi)$  le scalaire de  $\overline{\mathbb{K}}$  obtenu par substitution de  $x_i = \chi_i$  dans  $r$  pour tout  $i = 1, \dots, n$ . Nous avons alors le théorème suivant.

**Théorème 3.8.** (Théorème de Stickelberger)[Rou99] Supposons que  $\mathbb{K}$  est un corps de caractéristique zéro<sup>1</sup>. Soit  $r(x_1, \dots, x_n) \in R$ . Alors, les valeurs propres de  $M_r$  sont exactement les  $r(\chi)$ , où  $\chi \in V(I)$ , en comptant les multiplicités  $\mu_\chi$ .

Autrement dit, en calculant le polynôme caractéristique de  $M_r$ , on obtient un polynôme  $p_r$  dont les racines sont exactement  $r(\chi)$  où  $\chi$  est dans  $V(I)$ . Notons que si  $p_r$  est sans facteur carré, nous pouvons en conclure que  $r$  sépare  $V(I)$ .

Les matrices de multiplication et le Théorème de Stickelberger servent notamment au calcul de la *Représentation Univariée Rationnelle* présentée au Théorème 3.17.

### 3.1.9 Résultants

Soient deux polynômes  $f$  et  $g$  dans  $\mathbb{K}[x]$ , respectivement de degré  $m$  et  $n$  :

$$\begin{cases} f = a_m x^m + \dots + a_1 x + a_0, \\ g = b_n x^n + \dots + b_1 x + b_0, \end{cases} \quad a_0, \dots, a_m, b_0, \dots, b_n \in \mathbb{K}.$$

---

1. La caractéristique d'un anneau est le plus petit entier  $n \geq 1$  tel que  $\sum_{i=1}^n 1 = 0$ . Si un tel entier  $n$  n'existe pas, alors l'anneau est dit de caractéristique zéro. Cette définition s'étend aux corps.

Le *résultant* de  $f$  et  $g$  par rapport à la variable  $x$  est défini comme étant le déterminant de la *matrice de Sylvester* de taille  $(m+n) \times (m+n)$  suivante

$$\text{Syl}(f, g) = \begin{pmatrix} U \\ V \end{pmatrix} \in \mathbb{K}^{(m+n) \times (m+n)},$$

$$U = \begin{pmatrix} a_m & a_{m-1} & \dots & a_n & \dots & a_0 & 0 & \dots & 0 \\ 0 & a_m & \dots & a_{n-1} & \dots & a_1 & a_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_m & \dots & \dots & \dots & \dots & a_0 \end{pmatrix} \in \mathbb{K}^{(n) \times (m+n)},$$

$$V = \begin{pmatrix} b_n & b_{n-1} & \dots & b_m & \dots & b_0 & 0 & \dots & 0 \\ 0 & b_n & \dots & b_{m-1} & \dots & b_1 & b_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 & b_n & \dots & \dots & \dots & \dots & b_0 \end{pmatrix} \in \mathbb{K}^{(m) \times (m+n)},$$

c'est-à-dire  $\text{Res}(f, g, x) := \det(\text{Syl}(f, g)) \in \mathbb{K}$ .

Nous résumons ici quelques propriétés remarquables du résultant.

**Proposition 3.3.** [CLO06] Soient deux polynômes  $f$  et  $g$  de  $\mathbb{K}[x]$ . Alors,

1.  $\text{Res}(f, g, x)$  est un polynôme en les coefficients de  $f$  et  $g$ .
2.  $\text{Res}(f, g, x) = 0$  si et seulement si  $f$  et  $g$  ont un facteur commun dans  $\mathbb{K}[x]$ .
3. Il existe deux polynômes  $A$  et  $B$  dans  $\mathbb{K}[x]$  tels que :

$$\text{Res}(f, g, x) = A f + B g.$$

Les coefficients de  $A$  et  $B$  sont des polynômes en les coefficients de  $f$  et  $g$ .

Dans cette thèse, le point 2 de la précédente proposition sera souvent utilisé.

Notons aussi que la notion de résultant s'étend à un système de  $N \in \mathbb{N}$  polynômes. Pour plus de détails, nous référons à [CLO06, Chap. 3].

### 3.1.10 Variétés discriminantes

Supposons que l'on travaille dans le corps  $\mathbb{K} := \mathbb{Q}(\theta)$ , où  $\theta = (\theta_0, \dots, \theta_n)$  est un vecteur de paramètres, et que l'on étudie l'idéal zéro-dimensionnel  $\langle f_1, \dots, f_s \rangle$  de  $\mathbb{K}[x_1, \dots, x_n]$ . Le nombre de solutions (dans  $\overline{\mathbb{K}}$ , comptées avec multiplicités) du système polynomial  $\{f_1, \dots, f_s\}$  peut changer pour des valeurs différentes de  $\theta$ . Pour s'en convaincre, nous pouvons considérer le système polynomial suivant :

$$\begin{cases} f_1 := \theta x_1 - 1 = 0, \\ f_2 := x_2 - 1 = 0. \end{cases}$$

Pour  $\theta \neq 0$ , ce système admet une solution simple  $(1/\theta, 1)$ . En revanche, pour  $\theta = 0$ , ce système n'admet pas de solutions. Il convient alors de caractériser ces valeurs de paramètres

pour lesquelles le nombre de solutions d'un système polynomial change. C'est ce que caractérise une *variété discriminante*.

Définir précisément une *variété discriminante* d'une manière générale est technique. Nous proposons ici de limiter sa définition à un polynôme univarié de degré  $n$ . Pour une définition plus générale, nous référons à [LR07] ou [Mor08]. Notons alors  $\theta := (\theta_0, \dots, \theta_n)$  un vecteur de paramètres et définissons un polynôme comme suit :

$$P := \theta_n X^n + \theta_{n-1} X^{n-1} + \dots + \theta_1 X + \theta_0.$$

Tout d'abord, définissons la notion de *discriminant* de  $P$ .

**Définition 3.18.** Le *discriminant* de  $P$  est défini par

$$\Delta_P := \text{Res} \left( P, \frac{dP}{dx}, x \right), \quad (3.3)$$

où  $\text{Res} \left( P, \frac{dP}{dx}, x \right)$  dénote le résultant de  $P$  et de sa dérivée en  $x$ . Nous adoptons aussi la notation :

$$\text{disc}(P, x) := \text{Res} \left( P, \frac{dP}{dx}, x \right).$$

Pour plus d'informations sur les résultants, voir la Section 3.1.9.

En pratique, le discriminant de  $P$  correspond au lieu des paramètres où l'une des racines de  $P$  devient multiple. Des logiciels de calcul formel tels que `Maple` ou `Mathematica` permettent de calculer cet objet.

**Exemple 3.7.** Soit  $P := \theta_2 X^2 + \theta_1 X + \theta_0$ . Le discriminant de  $P$  est alors défini par :

$$\Delta_P := \theta_1^2 - 4\theta_2\theta_0.$$

$\Delta_P = 0$  est le lieu de l'espace des paramètres  $\theta$  sur lequel  $P$  admet une solution multiple.

Du discriminant de  $P$ , nous définissons sa variété discriminante.

**Définition 3.19.** La *variété discriminante* de  $P$  est définie par :

$$\mathcal{V}_P := \{ \theta \in \mathbb{C}^{n+1} \mid \Delta_P(\theta) = 0 \} \cup \{ \theta \in \mathbb{C}^{n+1} \mid \theta_n = 0 \}. \quad (3.4)$$

**Exemple 3.8.** Soit  $P := \theta_2 X^2 + \theta_1 X + \theta_0$ . Le variété discriminante de  $P$  est alors définie par :

$$\mathcal{V}_P := \{ \theta \in \mathbb{C}^3 \mid \theta_1^2 - 4\theta_2\theta_0 = 0 \} \cup \{ \theta \in \mathbb{C}^3 \mid \theta_2 = 0 \}.$$

Une coupe de  $\mathcal{V}_P$  à  $\theta_2 = 1$  est donnée en Figure 3.1.

D'une manière générale, la variété discriminante d'un système polynomial est une généralisation du discriminant d'un polynôme univarié. Elle décrit l'ensemble des points critiques du système, c'est-à-dire les singularités, le lieu où le système admet une solution de multiplicité plus grande que 1, et le lieu où une solution tend vers l'infini. Géométriquement, la variété discriminante permet de décomposer l'espace des paramètres en *cellules régulières* dans lesquelles le nombre de racines réelles est constant. En prenant un point particulier dans chaque cellule régulière, et en calculant les solutions du système en ce point (via, par exemple, un algorithme d'isolation certifié tel que celui proposé dans [CA76]), nous déduisons

alors le nombre de solutions réelles dans la cellule.

Bien que les frontières des cellules soient données par la variété  $\mathcal{V}_P$ , trouver un point témoin dans chacune d'elles n'est pas trivial. Il convient alors de définir une méthode systématique pour trouver un point dans chaque cellule. Une méthode consiste à calculer une *Décomposition Cylindrique Algébrique* [CH91, ACM84] de  $P$  et de ne considérer que les cellules de dimension maximale (c'est-à-dire de dimension égale au nombre de paramètres, ou en d'autres termes, en ne considérant que les cellules où les racines sont de multiplicité 1). Nous proposons d'introduire une telle méthode via l'exemple simple suivant.

**Exemple 3.9.** Considérons un polynôme  $P$  de la forme

$$P(b_1) := b_1^2 + \theta_1 b_1 + \theta_0.$$

Le coefficient de tête de  $P$  étant 1, sa variété discriminante est :

$$\mathcal{V}_P = \{(\theta_0, \theta_1) \in \mathbb{R}^2 \mid \Delta_P(\theta) := \theta_1^2 - 4\theta_0 = 0\}.$$

Celle-ci est représentée en Figure 3.1.

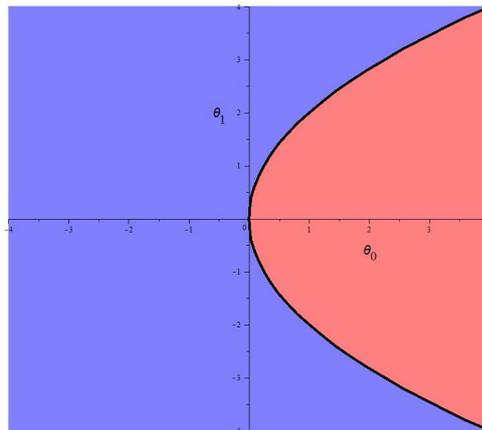


FIGURE 3.1 – Variété discriminante de  $P$  suivant  $b_1$

Nous distinguons alors deux cellules. Afin de les caractériser, l'idée est de calculer certains *points singuliers* de la variété  $\mathcal{V}_P$  [CLO15, Chap. 3], c'est-à-dire des points d'intersection de la variété avec elle-même, ou des points où une dérivée s'annule.

Considérons  $\Delta_P$  comme un polynôme en  $\theta_1$ . Nous avons :

$$\frac{\partial \Delta_P}{\partial \theta_1} = 2\theta_1.$$

Afin de déterminer les points de  $\Delta_P$  où sa dérivée en  $\theta_1$  s'annule, nous devons considérer le système polynomial suivant :

$$\begin{cases} \Delta_P(\theta_1) = \theta_1^2 - 4\theta_0 = 0, \\ \frac{\partial \Delta_P}{\partial \theta_1}(\theta_1) = 2\theta_1 = 0, \end{cases} \quad (3.5)$$

Ceci rappelle la définition du discriminant. En effet, nous cherchons  $\theta_1$  tel que celui-ci soit racine de  $\Delta_P$  et de sa dérivée en  $\theta_1$ . En d'autres termes, nous cherchons les conditions telles que  $\theta_1$  soit racine multiple de  $\Delta_P$ . Le calcul du discriminant de  $\Delta_P$  en  $\theta_1$  donne :

$$\text{disc}(\Delta_P, \theta_1) = 16 \theta_0. \quad (3.6)$$

Ce dernier s'annule si et seulement si  $\theta_0 = 0$ . En particulierisant  $\theta_0 = 0$  dans  $\Delta_P$ , nous obtenons  $\theta_1^2 = 0$ , soit  $\theta_1 = 0$ . Nous trouvons alors un seul point critique :  $\theta = (0, 0)$ . Notons que la résolution directe du système (3.5) nous aurait permis de trouver ce point directement. La Figure 3.1 nous permet de confirmer un tel résultat.

Notons alors que (3.6) décompose les valeurs de  $\theta_0$  en deux intervalles  $I_1$  et  $I_2$  :

$$I_1 := \{\theta_0 \in \mathbb{R} \mid \theta_0 < 0\}, \quad I_2 := \{\theta_0 \in \mathbb{R} \mid \theta_0 > 0\}. \quad (3.7)$$

Au dessus de chaque intervalle, le nombre de racines réelles de  $\Delta_P$  en  $\theta_1$  est constant. Plus particulièrement, si  $\theta_0 < 0$ , alors  $\Delta_P$  n'admet pas de racine réelle en  $\theta_1$ , et si  $\theta_0 > 0$ , alors  $\Delta_P$  admet 2 racines réelles en  $\theta_1$ . Nous pouvons alors décomposer l'espace  $\theta = (\theta_0, \theta_1)$  en 4 cellules notées  $C_i$ ,  $i = 1, \dots, 4$ , définies par :

- $C_1 := \{\theta = (\theta_0, \theta_1) \in \mathbb{R}^2 \mid \theta_0 < 0\}$ ,
- $C_2 := \{\theta = (\theta_0, \theta_1) \in \mathbb{R}^2 \mid \theta_0 > 0, \theta_1 < -\sqrt{4\theta_0}\}$ ,
- $C_3 := \{\theta = (\theta_0, \theta_1) \in \mathbb{R}^2 \mid \theta_0 > 0, -\sqrt{4\theta_0} < \theta_1 < \sqrt{4\theta_0}\}$ ,
- $C_4 := \{\theta = (\theta_0, \theta_1) \in \mathbb{R}^2 \mid \theta_0 > 0, \theta_1 > \sqrt{4\theta_0}\}$ .

Dans chaque cellule  $C_i$ , notons que, le nombre de racines réelles de  $P$  est constant puisque  $\Delta_P$  ne s'y annule pas. De plus, un point témoin dans chacune d'elle est facilement calculable comme suit. On choisit un  $\theta_0$  positif puis on particularise  $\Delta_P$  en ce point, dont on calcule les racines en  $\theta_1$ . Il suffit alors de choisir un point dans un intervalle défini par ces dernières racines en  $\theta_1$ .

Par ailleurs, notons qu'en choisissant de calculer le discriminant de  $\Delta_P$  en  $\theta_0$ , nous obtiendrions une décomposition en cellule plus simple puisque celui-ci ne s'annule jamais. En effet, pour tout  $\theta_1$ , nous avons :

$$\Delta_P(\theta_0, \theta_1) = 0 \quad \Leftrightarrow \quad \theta_0 = \theta_1^2/4.$$

En procédant de la sorte, nous obtenons alors uniquement deux cellules discriminées par les valeurs de  $\theta_0$  suivant les valeurs de  $\theta_1$ . Ceci est facilement remarquable en tournant la Figure 3.1 d'un quart de tour.

□

Une telle procédure se généralise à un nombre quelconque de paramètres en calculant des discriminants successifs de la variété discriminante du polynôme étudié. Les polynômes obtenus à ces différentes étapes sont appelés *polynômes projetés*.

De cet exemple, retenons simplement la chose suivante : en calculant les discriminants successifs de  $\Delta_P$  suivant les paramètres  $\theta$  dans un ordre donné, nous décomposons l'espace des paramètres en cellules dans lesquelles le nombre de racines réelles de  $P$  est constant, et dont il est possible d'extraire un point particulier facilement. En pratique, nous utilisons le système de calcul formel *Maple* et la commande *CellDecomposition* du package *RootFinding* afin de calculer une telle décomposition en cellules.

## 3.2 Classe de systèmes et équation de Riccati algébrique $\mathcal{R} = 0$ étudiés

Dans ce chapitre, nous considérons un système linéaire SISO  $G$ , de dimension finie  $n \geq 1$ , strictement propre, et représenté par sa fonction de transfert  $G$  d'entrée  $u$  et de sortie  $y$ . Soient les paramètres  $a$  et  $c$  suivants :

$$\begin{cases} a := (a_0, \dots, a_{n-1}), \\ c := (c_0, \dots, c_{n-1}). \end{cases}$$

Nous notons par  $\mathbb{K}[x]$  l'ensemble des polynômes en la variable  $x$  à coefficients dans un corps  $\mathbb{K}$ , et  $\mathbb{K}(x)$  l'ensemble des fractions rationnelles en  $x$  à coefficients dans un corps  $\mathbb{K}$ .

Soient les polynômes suivants

$$\begin{cases} D(s) := a_0 + a_1 s + \dots + a_i s^i + \dots + a_{n-1} s^{n-1} + s^n \in \mathbb{Q}[a, s], \\ N(s) := c_0 + c_1 s + \dots + c_j s^j + \dots + c_{n-1} s^{n-1} \in \mathbb{Q}[c, s], \end{cases}$$

tels que :

$$G(s) := \frac{N(s)}{D(s)} \in \mathbb{Q}(a, c, s). \quad (3.8)$$

Considérons la représentation d'état de  $G$  sous la forme *canonique de commande* définie par le triplet  $(A, B, C)$  donné comme suit :

$$\begin{cases} \dot{X} = A X + B u, \\ y = C X, \end{cases}$$

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix} \in \mathbb{Q}[a]^{n \times n}, \quad B := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{Q}^{n \times 1}, \quad (3.9)$$

$$C := (c_0 \quad c_1 \quad \dots \quad c_{n-1}) \in \mathbb{Q}[c]^{1 \times n}.$$

**Remarque 3.7.** Supposons que le système  $G$  dépende de paramètres physiques  $\theta$  (tels qu'une masse, une raideur, une longueur, etc.). Alors, les paramètres  $a$  et  $c$  dépendent explicitement de  $\theta$ .

Afin de simplifier les futures notations, nous imposons les notations suivantes :

$$\begin{cases} a_n := 1, \\ c_n := 0, \\ \forall k \in \mathbb{Z}, \quad k < 0, \quad a_k := 0, \quad c_k := 0, \\ \forall k \in \mathbb{Z}, \quad k > n, \quad a_k := 0, \quad c_k := 0. \end{cases} \quad (3.10)$$

Dans le cas des systèmes SISO, nous souhaitons résoudre l'ERA associée au PCR pour un tel système  $G$ , à savoir (voir Corollaire 2.6) :

$$\mathcal{R} := X A + A^T X - X B B^T X + C^T C = 0. \quad (3.11)$$

Une telle équation peut-être vue comme un système d'équations polynomiales en les coefficients de  $X$ , comme nous le verrons par la suite. Par résoudre, nous entendons trouver une représentation la plus simple des solutions de  $\mathcal{R} = 0$ . Dans ce chapitre, nous ferons référence au PCR dans le cas des systèmes SISO sans préciser davantage le contexte précédent.

### 3.3 Réduction de $\mathcal{R} = 0$ à un système d'équations polynomiales $\mathcal{B} = 0$

#### 3.3.1 Paramétrisation des solutions de $\mathcal{R} = 0$ par un calcul des sous-espaces propres de la matrice Hamiltonienne associée

A paramètres fixés  $a$  et  $c$ , une approche classique pour trouver la solution définie positive de  $\mathcal{R} = 0$  est de calculer les sous-espaces propres de la *matrice Hamiltonienne* associée [ACA<sup>+</sup>99]. Cette démarche fut reprise par [KYAH09] pour résoudre  $\mathcal{R} = 0$  avec des paramètres fixés. Nous proposons de reprendre cette démarche ici, dont nous réutiliserons des résultats par la suite.

#### Propriétés de la matrice Hamiltonienne associée à l'EAR

La *matrice Hamiltonienne* associée à l'équation de Riccati algébrique (3.11) est une matrice de taille  $2n \times 2n$  définie par :

$$\mathcal{H} := \begin{pmatrix} A & -B B^T \\ -C^T C & -A^T \end{pmatrix} \in \mathbb{Q}[a, c]^{2n \times 2n}. \quad (3.12)$$

Définissons son polynôme caractéristique comme suit.

**Définition 3.20.** Nous dénotons par  $f$  le polynôme caractéristique de  $\mathcal{H}$ , soit :

$$f(\lambda) := \det(\mathcal{H} - \lambda I_{2n}). \quad (3.13)$$

**Remarque 3.8.** Notons que le concept de déterminant est bien défini sur un anneau commutatif tel que  $\mathbb{Q}[a, c]$ . Voir, par exemple, [Rot09].

Ce polynôme possède la propriété suivante.

**Théorème 3.9.** Soient

$$d_{2k} := a_k^2 + c_k^2 + 2 \sum_{m=1}^n (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}), \quad (3.14)$$

Alors, le polynôme caractéristique de  $\mathcal{H}$  a la forme suivante :

$$f(\lambda) := (-1)^n \sum_{k=0}^n (-1)^k d_{2k} \lambda^{2k}. \quad (3.15)$$

*Démonstration.* Rappelons que le polynôme caractéristique de  $\mathcal{H}$  (voir (3.12)) est défini par :

$$f(\lambda) := \det(\mathcal{H} - \lambda I_{2n}) = (-1)^{2n} \det(\lambda I_{2n} - \mathcal{H}) = \det(\lambda I_{2n} - \mathcal{H}).$$

Plus précisément, ce déterminant s'écrit

$$f(\lambda) = \det(H^{(0)}),$$

où :

$$H^{(0)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & -1 & 0 & \dots & 0 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & -1 & 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \lambda & -1 & 0 & \dots & 0 & 0 \\ a_0 & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} + \lambda & 0 & \dots & 0 & 1 \\ \\ c_0^2 & c_1 c_0 & \dots & \dots & \dots & c_{n-1} c_0 & \lambda & 0 & 0 & \dots & 0 & -a_0 \\ c_0 c_1 & c_1^2 & \dots & \dots & \dots & c_{n-1} c_1 & 1 & \lambda & 0 & \dots & 0 & -a_1 \\ c_0 c_2 & c_1 c_2 & \dots & \dots & \dots & c_{n-1} c_2 & 0 & 1 & \lambda & \dots & 0 & -a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 & 1 & \lambda & -a_{n-2} \\ c_0 c_{n-1} & c_1 c_{n-1} & \dots & \dots & \dots & c_{n-1}^2 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} + \lambda \end{pmatrix}.$$

Rappelons quelques propriétés du déterminant que nous allons utiliser [Gri15, Chap. 4]. Le déterminant d'une matrice  $M$  à coefficients dans un anneau commutatif tel que  $\mathbb{Q}[a, c]$  possède les propriétés suivantes :

- Échanger une ligne de  $M$  avec une autre, ou une colonne avec une autre multiplie le déterminant par  $-1$ .
- La valeur du déterminant de  $M$  ne change pas lorsque l'on ajoute à une colonne de  $M$  une combinaison linéaire des autres colonnes de  $M$ . Cette propriété est aussi vérifiée pour les lignes de la matrice  $M$ .

Nous noterons par  $\text{ligne}_i(M)$  la  $i^{\text{ème}}$  ligne d'une matrice  $M$ , et par  $\text{col}_i(M)$  sa  $i^{\text{ème}}$  colonne. Nous allons alors effectuer des opérations sur les lignes et les colonnes de  $H^{(0)}$  afin de calculer son déterminant. Définissons la matrice  $H^{(1)} \in \mathbb{Q}[a, c, \lambda]^{2n \times 2n}$  comme suit :

$$H_{i,j}^{(1)} := \begin{cases} -\lambda H_{n+2,j}^{(0)} + H_{n+1,j}^{(0)} & \text{si } i = n + 1, \\ H_{i,j}^{(0)} & \text{sinon.} \end{cases}$$

La matrice  $H_1$  a alors la forme suivante :

$$H^{(1)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & \ddots & 0 & \dots & 0 & \vdots & \dots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & & \ddots & \lambda & -1 & 0 & \vdots & \dots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \lambda & -1 & 0 & \dots & 0 & 0 \\ a_0 & a_1 & \dots & \dots & a_{n-2} & a_{n-1} + \lambda & 0 & \dots & 0 & 1 \\ \\ c_0(c_0 - \lambda c_1) & c_1(c_0 - \lambda c_1) & \dots & \dots & \dots & c_{n-1}(c_0 - \lambda c_1) & 0 & -\lambda^2 & 0 & \dots & 0 & -a_0 + \lambda a_1 \\ c_0 c_1 & c_1^2 & \dots & \dots & \dots & c_{n-1} c_1 & 1 & \lambda & 0 & \dots & 0 & -a_1 \\ c_0 c_2 & c_1 c_2 & \dots & \dots & \dots & c_{n-1} c_2 & 0 & 1 & \lambda & \dots & 0 & -a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 & 1 & \lambda & -a_{n-2} \\ c_0 c_{n-1} & c_1 c_{n-1} & \dots & \dots & \dots & c_{n-1}^2 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} + \lambda \end{pmatrix}.$$

Alors, par construction, nous avons :

$$f(\lambda) = \det(H^{(1)}).$$

La  $(n+1)$ ème colonne de  $H^{(1)}$  est alors composée uniquement de zéros et d'un seul terme égal à 1 (le terme à la ligne  $n+2$ ). Cette colonne peut-être utilisée pour mettre à zéros tous les coefficients de la ligne  $n+2$ . Pour cela, nous utilisons la procédure suivante.

### Réduction 3.1.

$$H^{(2)} := H^{(1)}$$

$$i_0 := n + 2$$

$$j_0 := n + 1$$

for  $j$  from 1 to  $2n$

if  $j \neq j_0$  then

$$\text{col}_j(H^{(2)}) := \text{col}_j(H^{(1)}) - \frac{H_{i_0,j}^{(1)}}{H_{i_0,j_0}^{(1)}} \times \text{col}_{j_0}(H^{(1)})$$

end if

end for.

En appliquant la Réduction 3.1, nous créons la matrice

$$H^{(2)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & -1 & 0 & \dots & 0 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & -1 & 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \lambda & -1 & 0 & \dots & 0 & 0 \\ a_0 & a_1 & \dots & \dots & a_{n-2} & a_{n-1} + \lambda & 0 & \dots & 0 & 1 \\ \\ c_0(c_0 - \lambda c_1) & c_1(c_0 - \lambda c_1) & \dots & \dots & \dots & c_{n-1}(c_0 - \lambda c_1) & 0 & -\lambda^2 & 0 & \dots & 0 & -a_0 + \lambda a_1 \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ c_0 c_2 & c_1 c_2 & \dots & \dots & \dots & c_{n-1} c_2 & 0 & 1 & \lambda & \dots & 0 & -a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 & 1 & \lambda & -a_{n-2} \\ c_0 c_{n-1} & c_1 c_{n-1} & \dots & \dots & \dots & c_{n-1}^2 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} + \lambda \end{pmatrix},$$

telle que :

$$f(\lambda) = \det(H^{(2)}).$$

Ces deux dernières opérations peuvent être répétées de la ligne  $n + 1$  jusqu'à la ligne  $2n$ . Pour plus de clarté, nous explicitons la démarche dans la procédure suivante.

### Réduction 3.2.

```

 $H^{(3)} := H^{(2)}$ 
for  $i$  from  $n + 2$  to  $2n$ 
  ligne $_{n+1}(H^{(3)}) :=$  ligne $_{n+1}(H^{(3)}) +$  ligne $_i(H^{(3)}) \times (-1)^{i-(n+1)} \lambda^{i-(n+1)}$ 
  for  $j$  from 1 to  $2n$ 
    if  $j \neq i - 1$  then
      col $_j(H^{(3)}) :=$  col $_j(H^{(3)}) - H_{i,j}^{(3)} \times$  col $_{i-1}(H^{(3)})$ 
    end if
  end for
end for.
```

En appliquant la Réduction 3.2, et en utilisant le fait que  $a_n := 1$ , nous obtenons la

matrice suivante

$$H^{(3)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & \ddots & 0 & \dots & 0 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & -1 & 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \lambda & -1 & 0 & \dots & 0 & 0 \\ a_0 & \dots & a_j & \dots & a_{n-2} & a_{n-1} + \lambda & 0 & \dots & 0 & 1 \\ \dots & \dots & c_j \sum_{k=0}^{n-1} (-1)^k c_k \lambda^k & \dots & \dots & \dots & 0 & 0 & 0 & \dots & 0 & -\sum_{k=0}^n (-1)^k a_k \lambda^k \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 & 0 & \dots & \dots & 0 & 1 & 0 \end{pmatrix},$$

vérifiant  $f(\lambda) = \det(H^{(3)})$ . Définissons la matrice  $H^{(4)}$  de la manière suivante :

$$H_{i,j}^{(4)} := \begin{cases} H_{n+1,j}^{(3)} + \left( \sum_{k=0}^n (-1)^k a_k \lambda^k \right) \times H_{n,j}^{(3)} & \text{si } i = n + 1, \\ H_{i,j}^{(3)} & \text{sinon.} \end{cases}$$

En rappelant que  $c_n = 0$  et  $a_n = 1$  (voir (3.10)), nous avons

$$H_{n+1,j}^{(4)} = c_{j-1} \sum_{k=0}^{n-1} (-1)^k c_k \lambda^k + a_{j-1} \sum_{k=0}^n (-1)^k a_k \lambda^k = \sum_{k=0}^n (-1)^k (c_{j-1} c_k + a_{j-1} a_k) \lambda^k,$$

pour  $j = 1, \dots, n-1$ . Nous introduisons alors la notation suivante :

$$\alpha_{j,k} := c_j c_k + a_j a_k = \alpha_{k,j}, \quad j = 0, \dots, n, \quad k = 0, \dots, n. \quad (3.16)$$

Nous avons alors :

$$H_{n+1,j}^{(4)} = \sum_{k=0}^n (-1)^k \alpha_{j-1,k} \lambda^k, \quad j = 1, \dots, n-1.$$

Par ailleurs, le coefficient à la ligne  $n+1$  et à la colonne  $n$  de  $H^{(4)}$  s'écrit :

$$H_{n+1,n}^{(4)} = \sum_{k=0}^n (-1)^k \alpha_{n-1,k} \lambda^k + \lambda \sum_{k=0}^n (-1)^k a_k \lambda^k.$$

La matrice  $H^{(4)}$  a alors la forme suivante

$$H^{(4)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & \lambda & -1 & \ddots & & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \ddots & \lambda & -1 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \lambda & -1 & 0 & \dots & \dots & 0 & 0 \\ a_0 & \dots & \dots & \dots & a_{n-2} & a_{n-1} + \lambda & 0 & \dots & \dots & 0 & 1 \\ \sum_{k=0}^n (-1)^k \alpha_{0,k} \lambda^k & \dots & \sum_{k=0}^n (-1)^k \alpha_{j,k} \lambda^k & \dots & \sum_{k=0}^n (-1)^k \alpha_{n-2,k} \lambda^k & \sum_{k=0}^n (-1)^k \alpha_{n-1,k} \lambda^k & 0 & 0 & 0 & \dots & 0 \\ & & & & & + \lambda \sum_{k=0}^n (-1)^k a_k \lambda^k & & & & & \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix},$$

et vérifie  $f(\lambda) = \det(H^{(4)})$ . Nous pouvons alors développer le déterminant par rapport à la dernière colonne de  $H^{(4)}$ , constituée d'un seul 1 à la ligne  $n$  et de zéros ailleurs. Nous devons alors multiplier le déterminant par  $(-1)^{2n+n} = (-1)^n$ . Ainsi, le déterminant à calculer est celui d'une matrice de taille  $(2n-1) \times (2n-1)$  et s'écrit  $f(\lambda) = (-1)^n \det(H^{(5)})$  où :

$$H^{(5)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda & -1 & 0 & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \ddots & \lambda & -1 & 0 & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \lambda & -1 & 0 & \dots & 0 \\ \sum_{k=0}^n (-1)^k \alpha_{0,k} \lambda^k & \dots & \sum_{k=0}^n (-1)^k \alpha_{j,k} \lambda^k & \dots & \dots & \sum_{k=0}^n (-1)^k \alpha_{n-1,k} \lambda^k & 0 & 0 & 0 & \dots & 0 \\ & & & & & + \lambda \sum_{k=0}^n (-1)^k a_k \lambda^k & & & & & \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

En développant le déterminant par rapport aux  $n-1$  dernières colonnes de  $H^{(5)}$ , nous

obtenons  $f(\lambda) := (-1)^n \det(H^{(6)})$ , où :

$$H^{(6)} := \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & -1 & 0 \\ 0 & 0 & \dots & 0 & \lambda & -1 \\ \sum_{k=0}^n (-1)^k \alpha_{0,k} \lambda^k & \dots & \dots & \dots & \dots & \sum_{k=0}^n (-1)^k \alpha_{n-1,k} \lambda^k \\ & & & & & + \lambda \sum_{k=0}^n (-1)^k a_k \lambda^k \end{pmatrix}.$$

Afin de calculer plus facilement le déterminant de la matrice  $H^{(6)}$ , nous appliquons ensuite la procédure suivante.

### Réduction 3.3.

$$\begin{aligned} H^{(7)} &:= H^{(6)} \\ \text{for } i &\text{ from } 1 \text{ to } n-1 \\ \text{col}_{n-i}(H^{(7)}) &:= \text{col}_{n-i}(H^{(7)}) + \lambda \times \text{col}_{n-i+1}(H^{(7)}) \\ \text{end for} \\ \text{end for.} \end{aligned}$$

La matrice  $H^{(7)}$  obtenue a alors la forme suivante

$$H^{(7)} := \begin{pmatrix} 0 & -1 & 0 & \dots & \dots & 0 \\ 0 & 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & -1 & 0 \\ 0 & 0 & \dots & 0 & 0 & -1 \\ \sum_{l=0}^{n-1} \sum_{k=0}^n (-1)^k \alpha_{l,k} \lambda^{l+k} & \star & \dots & \dots & \dots & \star \\ + \sum_{k=0}^n (-1)^k a_k \lambda^{n+k} & & & & & \end{pmatrix},$$

où les étoiles représentent des termes qu'il n'est pas nécessaire de préciser pour le calcul de son déterminant, et nous avons  $f(\lambda) = (-1)^n \det(H^{(7)})$ . Les expressions des  $\alpha_{l,k}$  et  $a_n = 1$ ,  $c_n = 0$  permettent alors d'obtenir l'expression intermédiaire suivante :

$$\begin{aligned} f(\lambda) &:= (-1)^n (-1)^{n-1} (-1)^{n-1} \\ &\times \sum_{k=0}^n \left( \left( \sum_{l=0}^{n-1} (-1)^k (a_l a_k + c_l c_k) \lambda^{l+k} \right) + (-1)^k (a_n a_k + c_n c_k) \lambda^{n+k} \right) \\ &= (-1)^n \sum_{k=0}^n \sum_{l=0}^n (-1)^k (a_l a_k + c_l c_k) \lambda^{l+k}. \end{aligned} \tag{3.17}$$

Par ailleurs, nous avons :

$$\begin{aligned} f(\lambda) &= (-1)^n \sum_{k=0}^n \sum_{l=0}^n (-1)^k (a_l a_k + c_l c_k) \lambda^{l+k}, \\ &= (-1)^n \sum_{k=0}^n \left( \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k}} (-1)^i (a_i a_j + c_i c_j) \lambda^{2k} + \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k+1}} (-1)^i (a_i a_j + c_i c_j) \lambda^{2k+1} \right). \end{aligned}$$

D'une part, le coefficient en  $\lambda^{2k+1}$  de  $f$  s'écrit :

$$\begin{aligned} \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k+1}} (-1)^i (a_i a_j + c_i c_j) &= \sum_{l=0}^k (-1)^l (a_l a_{2k+1-l} + c_l c_{2k+1-l}) \\ &\quad + \sum_{l=0}^k (-1)^{2k+1-l} (a_{2k+1-l} a_l + c_{2k+1-l} c_l). \end{aligned}$$

Nous avons  $a_l a_{2k+1-l} + c_l c_{2k+1-l} = a_{2k+1-l} a_l + c_{2k+1-l} c_l$  et le signe de  $(-1)^l$  et de  $(-1)^{2k+1-l} = (-1)^{l+1}$  sont toujours opposés, ce qui implique

$$\sum_{i+j=2k+1} (-1)^i (a_i a_j + c_i c_j) = 0,$$

ce qui montre que  $f$  est un polynôme pair, et plus précisément que :

$$f(\lambda) = (-1)^n \sum_{k=0}^n \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k}} (-1)^i (a_i a_j + c_i c_j) \lambda^{2k}. \quad (3.18)$$

En gardant à l'esprit les notations (3.10), nous pouvons écrire :

$$\begin{aligned} &\sum_{\substack{0 \leq i, j \leq n \\ i+j=2k}} (-1)^i (a_i a_j + c_i c_j) \\ &= \sum_{i=0}^{2k} (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) \\ &\text{(les notations (3.10) permettent des indices négatifs ou supérieur à } n) \\ &= \sum_{i=0}^k (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) + \sum_{i=k+1}^{2k} (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) \\ &\text{(changement de variable : } l = 2k - i) \\ &= \sum_{i=0}^k (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) + \sum_{l=k-1}^0 (-1)^{2k-l} (a_{2k-l} a_l + c_{2k-l} c_l) \\ &\text{(la somme est commutative)} \\ &= \sum_{i=0}^k (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) + \sum_{l=0}^{k-1} (-1)^{2k-l} (a_{2k-l} a_l + c_{2k-l} c_l) \\ &= (-1)^k (a_k^2 + c_k^2) + \sum_{i=0}^{k-1} ((-1)^i + (-1)^{2k-i}) (a_i a_{2k-i} + c_i c_{2k-i}). \end{aligned}$$

En remarquant que

$$(-1)^i + (-1)^{2k-i} = (-1)^i + (-1)^i = 2(-1)^i,$$

nous pouvons simplifier l'expression précédente et nous obtenons :

$$\begin{aligned} \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k}} (-1)^i (a_i a_j + c_i c_j) &= (-1)^k (a_k^2 + c_k^2) + 2 \sum_{i=0}^{k-1} (-1)^i (a_i a_{2k-i} + c_i c_{2k-i}) \\ &\text{(changement de variable : } m = k - i) \\ &= (-1)^k (a_k^2 + c_k^2) + 2 \sum_{m=k}^1 (-1)^{k-m} (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}) \\ &\text{(la somme est commutative)} \\ &= (-1)^k (a_k^2 + c_k^2) + 2 \sum_{m=1}^k (-1)^{k-m} (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}) \\ &= (-1)^k \left( a_k^2 + c_k^2 + 2 \sum_{m=1}^k (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}) \right). \end{aligned}$$

L'équation (3.18) donne alors

$$\begin{aligned} f(\lambda) &= (-1)^n \sum_{k=0}^n \sum_{\substack{0 \leq i, j \leq n \\ i+j=2k}} (-1)^i (c_i c_j + a_i a_j) \lambda^{2k} \\ &= (-1)^n \sum_{k=0}^n (-1)^k \left( c_k^2 + a_k^2 + 2 \sum_{m=1}^k (-1)^m (c_{k-m} c_{k+m} + a_{k-m} a_{k+m}) \right) \lambda^{2k} \\ &= (-1)^n \sum_{k=0}^n (-1)^k d_{2k} \lambda^{2k}, \end{aligned}$$

ce qui conclut la preuve.  $\square$

**Remarque 3.9.** En particulier, le Théorème 3.9 montre que le polynôme  $f$  est *pair* quelles que soient les valeurs de paramètres  $a$  et  $c$ , c'est-à-dire :

$$f(\lambda) = f(-\lambda).$$

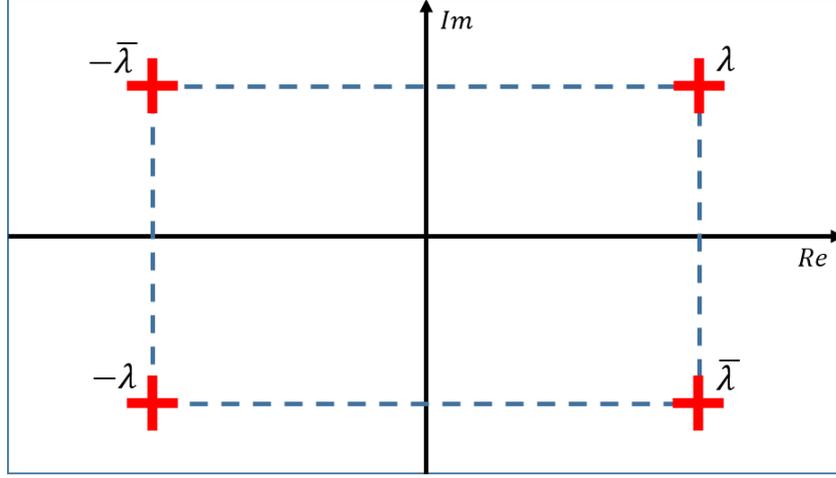
De plus, si  $a$  et  $c$  sont des paramètres réels, alors les  $d_{2k}$  définis par (3.14) sont réels, ce qui implique que le polynôme  $f$  est réel. Ainsi, si  $\lambda$  est racine de  $f$ , alors  $\bar{\lambda}$  l'est aussi. Nous avons alors les symétries  $\lambda \mapsto -\lambda$  et  $\lambda \mapsto \bar{\lambda}$  du spectre de la matrice  $\mathcal{H}$ .

Lorsque les paramètres  $a$  et  $c$  sont fixés à des valeurs réelles, nous avons le résultat suivant.

**Proposition 3.4.** [ACA<sup>+</sup>99] Soient  $a \in \mathbb{R}^n$  et  $c \in \mathbb{R}^n$ . Supposons que le système  $G$  défini en (3.8) est *stabilisable* et *déTECTABLE*. Alors la matrice Hamiltonienne  $\mathcal{H}$  définie en (3.12) n'a pas de valeurs propres sur l'axe imaginaire. De plus, toutes ses valeurs propres sont symétriques par rapport à l'axe imaginaire. Dans ce cas, les valeurs propres de  $\mathcal{H}$  ont alors la géométrie donnée par la Figure 3.2.

En d'autres termes, la Proposition 3.4 implique que, lorsque  $a$  et  $c$  sont fixés et que  $G$  est commandable et observable, il existe  $n$  valeurs propres de  $\mathcal{H}$ , notées  $\lambda_i$  pour  $i \in \llbracket 1, n \rrbracket$  et telles que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad i \neq j, \quad \lambda_i \neq -\lambda_j. \quad (3.19)$$


 FIGURE 3.2 – Schéma du spectre de  $\mathcal{H}$  avec  $\lambda \in \mathbb{C}_+$  vérifiant  $f(\lambda) = 0$ 

### Calcul d'une solution de $\mathcal{R} = 0$ à paramètres fixés

Soient  $a \in \mathbb{R}^n$  et  $c \in \mathbb{R}^n$ . Le raisonnement qui suit est basé sur l'article [KYAH09]. Supposons que le système est commandable et observable. Nous pouvons alors choisir  $n$  valeurs propres satisfaisant (3.19), que l'on note par  $\lambda_i$ ,  $i = 1, \dots, n$ . Soit  $V$  une matrice construite par concaténation verticale de vecteurs propres associée aux  $n$  valeurs propres  $\lambda_i$ . En notant  $v_{\lambda_i}$  un vecteur propre associé à la valeur propre  $\lambda_i$ , nous avons :

$$V := \begin{pmatrix} v_{\lambda_1} & v_{\lambda_2} & \cdots & v_{\lambda_n} \end{pmatrix} \in \mathbb{R}^{2n \times n}. \quad (3.20)$$

Nous avons alors le résultat suivant.

**Proposition 3.5.** Soient  $a \in \mathbb{R}^n$  et  $c \in \mathbb{R}^n$  tels que  $G$  soit commandable et observable. Soient  $\lambda_i$ ,  $i = 1, \dots, n$ ,  $n$  valeurs propres de  $\mathcal{H}$  satisfaisant (3.19). Soit  $V$  définie par (3.20) où  $v_{\lambda_i}$  est un vecteur propre associé à la valeur propre  $\lambda_i$ . Soit  $V_1$  et  $V_2$  deux matrices carrées de taille  $n$  vérifiant :

$$V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}.$$

Supposons que  $V_1$  soit inversible. La matrice  $X = V_2 V_1^{-1}$  est une solution de l'ERA (3.11).

*Démonstration.* Soit  $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$  la matrice diagonale contenant les  $n$  valeurs propres  $\lambda_i$  sur sa diagonale. Nous avons alors

$$\mathcal{H} V = V \Lambda, \quad (3.21)$$

soit

$$\begin{pmatrix} A & -B B^T \\ -C^T C & -A^T \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \Lambda,$$

c'est-à-dire :

$$\begin{cases} A V_1 - B B^T V_2 = V_1 \Lambda, \\ -C^T C V_1 - A^T V_2 = V_2 \Lambda. \end{cases}$$

Supposons  $V_1$  inversible. Notons alors  $X := V_2 V_1^{-1}$ . Nous avons

$$-C^T C = A^T X + V_2 \Lambda V_1^{-1},$$

et

$$\Lambda = V_1^{-1} A V_1 - V_1^{-1} B B^T V_2.$$

En combinant les deux précédentes équations, nous obtenons :

$$X A + A^T X - X B B^T X + C^T C = 0.$$

Alors,  $X$  est donc solution de  $\mathcal{R} = 0$ . □

**Remarque 3.10.** Une solution  $X$  calculée en utilisant la Proposition 3.5 n'est pas nécessairement définie positive.

La Proposition 3.5 montre que trouver une solution de  $\mathcal{R} = 0$  revient à un problème de calcul de vecteurs propres. Dans la prochaine section, nous montrons comment réutiliser ces résultats lorsque  $a$  et  $c$  ne sont pas fixés.

### Calcul d'une solution de $\mathcal{R} = 0$ à paramètres non-fixés par factorisation spectrale

Nous proposons maintenant d'étendre la méthode proposée dans [KYAH09] et expliquée dans la section précédente au cas où  $a$  et  $c$  ne sont pas fixés à des valeurs numériques. Montrons tout d'abord comment construire un vecteur propre de  $\mathcal{H}$ . Pour cela, notons  $\mathbb{K} := \mathbb{Q}(a, c)$  et définissons le polynôme  $q \in \mathbb{K}[x, y]$  vérifiant :

$$f(x) = (x - y) q(x, y) + f(y). \quad (3.22)$$

**Remarque 3.11.** Notons que  $q$  est bien un polynôme. Ceci peut se montrer en considérant le développement limité d'ordre  $N$  de  $f$  en  $y$  :

$$f(x) = f(y) + f'(y)(x - y) + \dots + \frac{f^{(N)}(y)}{N!} (x - y)^N + O((x - y)^N).$$

Pour  $N > \deg(f)$ , notons que le terme en  $O((x - y)^N)$  disparaît de l'expression précédente. Nous obtenons alors :

$$\frac{f(x) - f(y)}{x - y} = f'(y) + f''(y)(x - y) + \dots + \frac{f^{(N)}(y)}{N!} (x - y)^{N-1} + O((x - y)^{N-1}).$$

Pour  $N - 1 > \deg(f)$ , c'est-à-dire  $N > \deg(f) + 1$ , la partie droite de l'équation précédente étant polynomiale en  $x$  et  $y$ , alors  $q(x, y) \in \mathbb{K}[x, y]$ .

**Remarque 3.12.** Par ailleurs, un tel polynôme  $q$  peut se calculer en effectuant la division euclidienne de  $f(x)$  par  $x - y$  :

$$f(x) = (x - y) q(x, y) + f(y).$$

Alors, nous avons la proposition suivante.

**Proposition 3.6.** Soit l'anneau quotient  $\mathbb{A} = \mathbb{K}[\lambda]/\langle f(\lambda) \rangle$  et notons par  $\bar{\lambda}$  la classe de  $\lambda$  dans  $\mathbb{A}$ . Alors, nous avons :

$$\forall v \in \mathbb{A}^{2n}, \quad \mathcal{H} (q(\mathcal{H}, \bar{\lambda} I_{2n}) v) = \bar{\lambda} (q(\mathcal{H}, \bar{\lambda} I_{2n}) v). \quad (3.23)$$

Si  $f$  est un polynôme irréductible sur  $\mathbb{K}$ , alors  $\mathbb{A}$  est un corps et (3.23) montre que  $q(\mathcal{H}, \bar{\lambda} I_{2n}) v$  est un vecteur propre de  $\mathcal{H}$  pour la valeur propre  $\bar{\lambda} \in \mathbb{A}$ .

*Démonstration.* En substituant  $x = \mathcal{H}$  et  $y = \lambda I_{2n}$ , où  $\lambda$  est un paramètre, dans (3.22), et en utilisant  $f(\lambda I_{2n}) = f(\lambda) I_{2n}$  et le fait que le polynôme caractéristique  $f$  de  $\mathcal{H}$  vérifie  $f(\mathcal{H}) = 0$  [Rot09], nous obtenons :

$$(\mathcal{H} - \lambda I_{2n}) q(\mathcal{H}, \lambda I_{2n}) + f(\lambda) I_{2n} = 0. \quad (3.24)$$

L'identité (3.24) se réduit alors sur  $\mathbb{A}^{2n \times 2n}$  à  $(\mathcal{H} - \bar{\lambda} I_{2n}) q(\mathcal{H}, \bar{\lambda} I_{2n}) = 0$ , ce qui donne (3.23).  $\square$

Nous disposons alors d'un moyen de construire un vecteur propre de  $\mathcal{H}$  associé à n'importe quelle valeur propre de  $\mathcal{H}$ . Plutôt que de calculer directement les valeurs propres de  $\mathcal{H}$ , afin d'en déduire un vecteur propre, on peut considérer le problème suivant.

**Définition 3.21.** Rappelons que le polynôme caractéristique de  $\mathcal{H}$  s'écrit

$$f(\lambda) = (-1)^n \sum_{i=0}^n (-1)^i d_{2i} \lambda^{2i}, \quad (3.25)$$

où les  $d_{2i}$  sont définis en (3.14). Soit  $\beta := (\beta_0, \dots, \beta_{n-1})$  un vecteur de paramètres. Par convention, notons :

$$\begin{cases} \beta_n := 1, \\ \forall k \in \mathbb{Z}, k \leq 0, \beta_k = 0, \\ \forall k \in \mathbb{Z}, k > n, \beta_k = 0. \end{cases} \quad (3.26)$$

De plus, soit le polynôme

$$g(\lambda) := \sum_{i=1}^n \beta_i \lambda^i, \quad (3.27)$$

tels que :

$$f(\lambda) = (-1)^n g(\lambda) g(-\lambda). \quad (3.28)$$

La recherche d'un tel polynôme  $g$  est appelé *problème de factorisation de Graeffe*, en référence à la méthode de Graeffe [Hou59]. Lorsque  $g$  a toutes ses racines dans  $\mathbb{C}_-$ , nous parlerons de *problème de factorisation spectrale* [JK85].

En égalisant les coefficients en  $\lambda$  à gauche et à droite de (3.28), nous obtenons un système d'équations polynomiales que doivent satisfaire les  $\beta_k$ .

**Proposition 3.7.** [AHKY09] Les coefficients du polynôme  $g$  défini par (3.27), c'est-à-dire les  $\beta_k$ ,  $k = 0, \dots, n-1$ , satisfont le système polynomial suivant :

$$\left\{ \mathcal{B}_k := \beta_k^2 + 2 \sum_{m=1}^n (-1)^m \beta_{k-m} \beta_{k+m} - d_{2k} = 0 \quad | \quad k \in \llbracket 0, n-1 \rrbracket \right\}. \quad (3.29)$$

Ce système polynomial encode toutes les configurations possibles de (3.28). Nous notons alors  $\mathcal{B} := \{\mathcal{B}_k\}_{k=0, \dots, n-1}$ .

*Démonstration.* Le terme de droite de (3.28) s'écrit de la manière suivante :

$$(-1)^n g(\lambda) g(-\lambda) = (-1)^n \left( \sum_{k=0}^n \beta_k \lambda^k \right) \left( \sum_{l=0}^n \beta_l (-\lambda)^l \right) = (-1)^n \sum_{l=0}^n \sum_{k=0}^l (-1)^l \beta_k \beta_l \lambda^{k+l}.$$

Similairement au début du Théorème 3.9, nous pouvons prouver que les termes en  $\lambda^{2k+1}$  sont tous nuls et que nous avons :

$$(-1)^n g(\lambda) g(-\lambda) = (-1)^n \sum_{k=0}^n \sum_{i+j=2k} (-1)^i \beta_i \beta_j \lambda^{2k}.$$

En utilisant les conventions (3.26), nous obtenons alors :

$$f(\lambda) = (-1)^n \sum_{k=0}^n (-1)^k \left( \beta_k^2 + 2 \sum_{m=1}^n (-1)^m \beta_{k-m} \beta_{k+m} \right) \lambda^{2k}.$$

En égalisant les coefficients en  $\lambda$  à gauche et à droite de (3.28) et en utilisant (3.15), nous pouvons identifier les coefficients de  $f$  et montrer que

$$\forall k \in \mathbb{Z}, \quad \beta_k^2 + 2 \sum_{m=1}^n (-1)^m \beta_{k-m} \beta_{k+m} = d_{2k},$$

ce qui permet de conclure la preuve.  $\square$

Disposant de  $g$  satisfaisant (3.28), nous avons alors le résultat suivant.

**Théorème 3.10.** Avec les notations de la Proposition 3.6, soit  $\mathbb{L} := \mathbb{K}[\beta_0, \dots, \beta_{n-1}] / \langle \mathcal{B} \rangle$  tel que  $g = \sum_{i=0}^n \beta_i \lambda^i \in \mathbb{L}[\lambda]$ , où  $\beta_n = 1$ , satisfait une factorisation de Graeffe de  $f$ , c'est-à-dire vérifie  $f(\lambda) = (-1)^n g(\lambda) g(-\lambda)$ . De plus, notons  $\mathcal{C}$  la matrice suivante :

$$\mathcal{C} := \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \\ -\beta_0 & -\beta_1 & \dots & \dots & -\beta_{n-2} & -\beta_{n-1} \end{pmatrix} \in \mathbb{L}^{2n \times n}.$$

Alors, il existe une matrice  $\tilde{V} \in \mathbb{L}^{2n \times n}$  telle que :

$$\mathcal{H} \tilde{V} = \tilde{V} \mathcal{C}.$$

Finalement, avec les notations suivantes

$$\tilde{V} = \begin{pmatrix} \tilde{V}_1 \\ \tilde{V}_2 \end{pmatrix}, \quad \tilde{V}_1, \tilde{V}_2 \in \mathbb{L}^{n \times n},$$

si  $\tilde{V}_1$  est inversible, alors  $X = \tilde{V}_2 \tilde{V}_1^{-1}$  est solution de l'ERA associée à  $\mathcal{H}$ .

*Démonstration.* Notons  $\mathbb{B} := \mathbb{L}[\lambda] / \langle g(\lambda) \rangle = \mathbb{K}[\beta_0, \dots, \beta_{n-1}, \lambda] / \langle \mathcal{B}, g(\lambda) \rangle$  et  $\tilde{\lambda}$  la classe  $\lambda$  dans  $\mathbb{B}$ . L'identité (3.24) se réduit sur  $\mathbb{B}^{2n \times 2n}$  à  $(\mathcal{H} - \tilde{\lambda} I_{2n}) q(\mathcal{H}, \tilde{\lambda} I_{2n}) = 0$ , ce qui implique :

$$\forall \tilde{v} \in \mathbb{B}^{2n}, \quad \mathcal{H} \left( q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v} \right) = \tilde{\lambda} \left( q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v} \right). \quad (3.30)$$

En utilisant que  $\mathbb{B}$  est un  $\mathbb{L}$ -module libre de rang  $n$  [Rot09] et  $\{\tilde{\lambda}^i\}_{i=0,\dots,n-1}$  est une base de  $\mathbb{B}$ , nous pouvons écrire le vecteur  $q\left(\mathcal{H}, \tilde{\lambda} I_{2n}\right) \tilde{v} \in \mathbb{B}^{2n}$  dans la base précédente pour obtenir :

$$q\left(\mathcal{H}, \tilde{\lambda} I_{2n}\right) \tilde{v} = \tilde{V} \begin{pmatrix} 1 \\ \tilde{\lambda} \\ \vdots \\ \tilde{\lambda}^{n-1} \end{pmatrix}, \quad \tilde{V} \in \mathbb{L}^{2n \times n}.$$

L'équation (3.30) devient alors :

$$\begin{aligned} \mathcal{H} \tilde{V} \begin{pmatrix} 1 \\ \tilde{\lambda} \\ \vdots \\ \tilde{\lambda}^{n-1} \end{pmatrix} &= \tilde{\lambda} \tilde{V} \begin{pmatrix} 1 \\ \tilde{\lambda} \\ \vdots \\ \tilde{\lambda}^{n-1} \end{pmatrix} = \tilde{V} \begin{pmatrix} \tilde{\lambda} \\ \vdots \\ \tilde{\lambda}^n \end{pmatrix} \\ &= \tilde{V} \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \\ -\beta_0 & -\beta_1 & \dots & \dots & -\beta_{n-2} & -\beta_{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ \tilde{\lambda} \\ \vdots \\ \tilde{\lambda}^{n-1} \end{pmatrix}. \end{aligned}$$

Soit  $\mathcal{C} \in \mathbb{L}^{n \times n}$  la matrice compagnon précédente. Dans les bases canoniques, l'identité précédente donne alors :

$$\mathcal{H} \tilde{V} = \tilde{V} \mathcal{C}.$$

L'équation précédente se réécrit de la manière suivante

$$\begin{pmatrix} A & -B B^T \\ -C^T C & -A^T \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \mathcal{C},$$

c'est-à-dire :

$$\begin{cases} A V_1 - B B^T V_2 = V_1 \mathcal{C}, \\ -C^T C V_1 - A^T V_2 = V_2 \mathcal{C}. \end{cases}$$

Supposons  $V_1$  inversible. Notons alors  $X := V_2 V_1^{-1}$ . Nous avons

$$-C^T C = A^T X + V_2 \mathcal{C} V_1^{-1},$$

et

$$\mathcal{C} = V_1^{-1} A V_1 - V_1^{-1} B B^T V_2.$$

En combinant les deux précédentes équations, nous obtenons :

$$X A + A^T X - X B B^T X + C^T C = 0.$$

Alors,  $X$  est donc solution de  $\mathcal{R} = 0$ . □

**Remarque 3.13.** Le Théorème 3.10 montre que la méthode proposée dans [KYAH09] pour la paramétrisation des solutions de  $\mathcal{R} = 0$  à paramètres fixés peut s'étendre à paramètres non fixés par des considérations d'algèbre linéaire et de théorie des modules [Rot09]. De plus, cette méthode peut s'appliquer à des systèmes MIMO, car celle-ci ne nécessite que la connaissance de la matrice Hamiltonienne  $\mathcal{H}$ .

Résumons cette méthode à l'aide de l'algorithme suivant.

**Algorithme 3.3.**

Entrée : Représentation d'état  $(A, B, C)$  d'un système  $G$

Sortie : Matrice  $X$  solution de  $\mathcal{R} = 0$  dépendant de  $\beta := (\beta_0, \dots, \beta_{n-1})$  solution de  $\mathcal{B}$

1. Calculer

$$\mathcal{H} = \begin{pmatrix} A & -B B^T \\ -C^T C & -A^T \end{pmatrix}$$

et son polynôme caractéristique  $f(\lambda) = \sum_{k=0}^n d_{2k} \lambda^{2k}$ . Nous obtenons alors une expression des coefficients  $d_{2k}$  en fonction des coefficients des matrices  $A$ ,  $B$  et  $C$ .

2. Considérer le polynôme  $g(\lambda) = \sum_{k=0}^n \beta_k \lambda^k$  tel que :

$$f(\lambda) = (-1)^n g(\lambda) g(-\lambda).$$

3. Égaliser les coefficients en  $\lambda$  de  $f(\lambda)$  et de  $(-1)^n g(\lambda) g(-\lambda)$  afin d'obtenir le système polynomial  $\mathcal{B}$  suivant :

$$\left\{ \mathcal{B}_k := \beta_k^2 + 2 \sum_{m=1}^n (-1)^m \beta_{k-m} \beta_{k+m} = d_{2k} \quad | \quad k \in \llbracket 0, n-1 \rrbracket \right\}.$$

4. Calculer la division euclidienne de  $f(x)$  par  $(x - y)$  afin de construire le polynôme  $q$  tel que :

$$f(x) = (x - y) q(x, y) + f(y).$$

5. Calculer la matrice  $q(\mathcal{H}, \tilde{\lambda} I_{2n})$  en substituant  $x$  par  $\mathcal{H}$  et  $y$  par  $\tilde{\lambda} I_{2n}$  dans  $q(x, y)$ .
6. Choisir un vecteur  $\tilde{v} \in \mathbb{B}^{2n}$ , où  $\mathbb{B} := \mathbb{L}[\lambda] / \langle g(\lambda) \rangle$  avec  $\mathbb{L} := \mathbb{K}[\beta_0, \dots, \beta_{n-1}] / \langle \mathcal{B} \rangle$ , c'est-à-dire un vecteur de taille  $2n$  à coefficients polynomiaux en  $\tilde{\lambda}$  de degré au plus  $n-1$  et à coefficients dans  $\mathbb{L}$ . Nous obtenons alors un vecteur  $q(\mathcal{H}, \tilde{\lambda}) \tilde{v} \in \mathbb{B}^{2n}$ .
7. Calculer une forme normale des coefficients du vecteur  $q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v}$  par rapport à une base de Gröbner de l'idéal  $\langle g, \mathcal{B} \rangle$  afin d'obtenir le vecteur  $q(\mathcal{H}, \tilde{\lambda}) \tilde{v}$  dans la base  $\{1, \lambda, \dots, \tilde{\lambda}^{n-1}\}$  de  $\mathbb{B}$ . Il résulte une matrice  $\tilde{V}$  telle que :

$$q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v} = \tilde{V} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \tilde{\lambda}^{n-1} \end{pmatrix}.$$

8. Partitionner  $\tilde{V}$  en deux matrices de taille  $n \times n$  telles que :

$$\tilde{V} = \begin{pmatrix} \tilde{V}_1 \\ \tilde{V}_2 \end{pmatrix}.$$

9. Si  $V_1$  est non singulière, alors construire la matrice suivante :

$$X := \tilde{V}_2 \tilde{V}_1^{-1}.$$

Illustrons les résultats précédents à l'aide d'un exemple explicite.

**Exemple 3.10.** Soit un système masse-ressort décrit par la Figure 3.3, notamment étudié dans [ZDG96, §10.2]. Une masse  $m$  est reliée à un support fixe par un ressort de raideur  $k$  et un amortisseur d'amplitude  $\xi$ . Nous étudions le déplacement de la masse  $m$ , noté par  $y_1$ , lorsque la masse  $m$  est excitée par une force  $e_1$ .

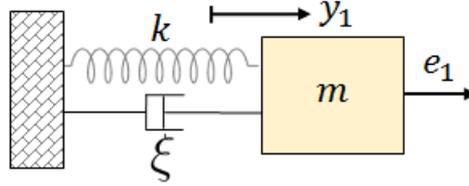


FIGURE 3.3 – Système masse-ressort encastré

La fonction de transfert du système à étudier est :

$$G := \frac{y_1}{e_1} = \frac{c_0}{s^2 + a_1 s + a_0}, \quad c_0 = \frac{1}{m}, \quad a_1 = \frac{\xi}{m}, \quad a_0 = \frac{k}{m}.$$

Il s'agit d'un système d'ordre 2. La représentation d'état de  $G$  sous forme canonique de commande est alors définie par :

$$A = \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = (c_0 \quad 0).$$

Notons  $\mathbb{K} := \mathbb{Q}(a_0, a_1, c_0)$ . La matrice Hamiltonienne  $\mathcal{H}$  s'écrit :

$$\mathcal{H} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -a_0 & -a_1 & 0 & -1 \\ -c_0^2 & 0 & 0 & a_0 \\ 0 & 0 & -1 & a_1 \end{pmatrix}.$$

Son polynôme caractéristique est donné par :

$$f(\lambda) = \lambda^4 + (-a_1^2 + 2a_0)\lambda^2 + a_0^2 + c_0^2.$$

Trouver une factorisation de Graeffe de  $\mathcal{H}$  consiste alors à trouver un polynôme de degré deux

$$g(\lambda) = \lambda^2 + \beta_1 \lambda + \beta_0,$$

tel que :

$$f(\lambda) = g(\lambda)g(-\lambda).$$

En égalisant les coefficients en  $\lambda$  de  $f(\lambda)$  et de  $g(\lambda)g(-\lambda)$ , nous obtenons que les coefficients de  $g$  vérifient le système polynomial suivant :

$$\begin{cases} \mathcal{B}_1 := \beta_1^2 - 2\beta_0 - a_1^2 + 2a_0 = 0, \\ \mathcal{B}_0 := \beta_0 - a_0^2 - c_0^2 = 0. \end{cases} \quad (3.31)$$

Notons alors  $\mathcal{B} := \{\mathcal{B}_0, \mathcal{B}_1\}$ ,  $\mathbb{L} := \mathbb{K}[\beta_0, \beta_1]/\langle \mathcal{B} \rangle$  et  $\mathbb{B} := \mathbb{L}[\lambda]/\langle g(\lambda) \rangle$ . Le polynôme  $q$ , défini en (3.22), s'écrit alors :

$$q(x, y) := x^3 + y x^2 + (y^2 - a_1^2 + 2 a_0) x + y^3 - y a_1^2 + 2 y a_0.$$

Notons  $\tilde{\lambda}$  la classe de  $\lambda$  dans  $\mathbb{B}$ . La substitution de  $x$  par  $\mathcal{H}$  et de  $y$  par  $\tilde{\lambda} I_{2n}$  dans  $q(x, y)$  donne alors :

$$q(\mathcal{H}, \tilde{\lambda} I_{2n}) = \begin{pmatrix} (\tilde{\lambda} + a_1)(\tilde{\lambda}^2 - \tilde{\lambda} a_1 + a_0) & \tilde{\lambda}^2 - \tilde{\lambda} a_1 + a_0 & 1 & -\tilde{\lambda} \\ -a_0^2 - \tilde{\lambda}(\tilde{\lambda} - a_1)a_0 - c_0^2 & \tilde{\lambda}(\tilde{\lambda}^2 - \tilde{\lambda} a_1 + a_0) & \tilde{\lambda} & -\tilde{\lambda}^2 \\ -\tilde{\lambda}^2 c_0^2 + a_1^2 c_0^2 & -c_0^2(\tilde{\lambda} - a_1) & (\tilde{\lambda} - a_1)(\tilde{\lambda}^2 + \tilde{\lambda} a_1 + a_0) & a_0^2 + \tilde{\lambda}(\tilde{\lambda} + a_1)a_0 + c_0^2 \\ c_0^2(\tilde{\lambda} + a_1) & c_0^2 & -\tilde{\lambda}^2 - \tilde{\lambda} a_1 - a_0 & \tilde{\lambda}(\tilde{\lambda}^2 + \tilde{\lambda} a_1 + a_0) \end{pmatrix}.$$

Le Théorème 3.10 montre que, pour tout  $\tilde{v} \in \mathbb{B}^{2n}$ ,  $q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v}$  est un "vecteur propre" de  $\mathcal{H}$  pour la valeur propre  $\tilde{\lambda} \in \mathbb{B}$ . En particulier, en choisissant le vecteur

$$\tilde{v} := (0 \quad 0 \quad 1 \quad 0)^T$$

nous obtenons alors :

$$q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v} = \begin{pmatrix} 1 \\ \tilde{\lambda} \\ (\tilde{\lambda} - a_1)(\tilde{\lambda}^2 + \tilde{\lambda} a_1 + a_0) \\ -\tilde{\lambda}^2 - \tilde{\lambda} a_1 - a_0 \end{pmatrix}.$$

Une base de  $\mathbb{B}$  est  $\{1, \tilde{\lambda}\}$ . Nous pouvons écrire le vecteur  $q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v}$  dans cette base pour obtenir

$$q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v} = \tilde{V} \begin{pmatrix} 1 \\ \tilde{\lambda} \end{pmatrix},$$

où :

$$\tilde{V} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -a_1 a_0 + \beta_0 \beta_1 & -a_0 + \beta_0 \\ -a_0 + \beta_0 & -a_1 + \beta_1 \end{pmatrix}.$$

La matrice  $\tilde{V}$  peut s'obtenir en calculant une forme normale des coefficients du vecteur  $q(\mathcal{H}, \tilde{\lambda} I_{2n}) \tilde{v}$  par rapport à une base de Gröbner de l'idéal suivant :

$$I := \langle g, \mathcal{B} \rangle = \langle \lambda^2 + \beta_1 \lambda + \beta_0, \beta_1^2 - 2 \beta_0 - a_1^2 + 2 a_0, \beta_0 - a_0^2 - c_0^2 \rangle.$$

Nous partitionnons alors  $\tilde{V}$  en deux matrices de taille  $2 \times 2$  telles que

$$\tilde{V} = \begin{pmatrix} \tilde{V}_1 \\ \tilde{V}_2 \end{pmatrix},$$

où :

$$\tilde{V}_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{V}_2 := \begin{pmatrix} -a_1 a_0 + \beta_0 \beta_1 & -a_0 + \beta_0 \\ -a_0 + \beta_0 & -a_1 + \beta_1 \end{pmatrix}.$$

Les solutions de  $\mathcal{R} = 0$  en fonction de  $\beta$  sont alors données par :

$$X := \tilde{V}_2 \tilde{V}_1^{-1} = \begin{pmatrix} -a_1 a_0 + \beta_0 \beta_1 & -a_0 + \beta_0 \\ -a_0 + \beta_0 & -a_1 + \beta_1 \end{pmatrix}.$$

**Remarque 3.14.** D'après (3.8), nous avons  $G = N/D$  où  $D(s) = \det(s I_n - A)$  de part la définition de  $G$  (voir la section précédente). De plus, notons que par une décomposition matricielle et les formules d'inversion matricielle ([ZDG96] p.23), nous avons :

$$\begin{aligned} \det(s I_{2n} - \mathcal{H}) &= \det(s I_n - A) \det((s I_n + A^T) - C^T C (s I_n - A)^{-1} B B^T), \\ &= D(s) \det(s I_n + A^T) \det(I_n - (s I_n + A^T)^{-1} C^T C (s I_n - A)^{-1} B B^T), \\ &= (-1)^n D(s) D(-s) \det(I_n - (s I_n + A^T)^{-1} C^T (C (s I_n - A)^{-1} B) B^T), \\ &= (-1)^n D(s) D(-s) \det(I_n - (s I_n + A^T)^{-1} C^T G(s) B^T). \end{aligned}$$

Or, pour  $M_1 \in \mathbb{R}^{q \times p}$  (resp.  $M_2 \in \mathbb{R}^{p \times q}$ ), nous avons

$$\det \begin{pmatrix} I_q & M_1 \\ -M_2 & I_p \end{pmatrix} = \det(I_p + M_1 M_2) = \det(I_q + M_2 M_1),$$

d'où :

$$\det(s I_{2n} - \mathcal{H}) = (-1)^n D(s) D(-s) \det(1 - B^T (s I_n + A^T)^{-1} C^T G(s)).$$

Sachant que  $G(-s) = G(-s)^T = B^T (-s I_n - A^T)^{-1} C^T$ , nous en déduisons :

$$\det(s I_{2n} - \mathcal{H}) = (-1)^n D(s) D(-s) (1 + G(s) G(-s)).$$

En développant  $G$ , on obtient alors :

$$\det(s I_{2n} - \mathcal{H}) = (-1)^n (D(s) D(-s) + N(s) N(-s)).$$

Le polynôme caractéristique de  $\mathcal{H}$  peut alors être également calculé à l'aide de l'expression de  $G$  :

$$f(\lambda) = (-1)^n (D(\lambda) D(-\lambda) + N(\lambda) N(-\lambda)). \quad (3.32)$$

On retrouve alors l'équation d'Euler-Lagrange découlant de la minimisation d'un critère quadratique portant sur l'énergie de l'entrée  $u$  et de la sortie  $y$  du système SISO  $G$ , c'est-à-dire :

$$\inf \int_0^\infty (y(t)^2 + u(t)^2) dt, \quad D \left( \frac{d}{dt} \right) y(t) = N \left( \frac{d}{dt} \right) u(t).$$

Pour plus de détails, voir [PQ04].

**Proposition 3.8.** Supposons que les paramètres  $a$  et  $c$  sont fixés à des valeurs numériques, c'est-à-dire que  $a, c \in \mathbb{R}^n$ . Soit  $g$  une factorisation spectrale de  $f$ , c'est-à-dire telle que :

$$\begin{cases} g(\lambda) = 0 \Leftrightarrow \lambda \in \mathbb{C}_-, \\ f(\lambda) = (-1)^n g(\lambda) g(-\lambda). \end{cases}$$

Alors, d'après (3.32), en posant

$$\widetilde{M}(s) := \frac{D(s)}{g(s)}, \quad \widetilde{M}^*(s) = \widetilde{M}(-s) = \frac{D(-s)}{g(-s)},$$

$$\widetilde{N}(s) := \frac{N(s)}{g(s)}, \quad \widetilde{N}^*(s) = \widetilde{N}(-s) = \frac{N(-s)}{g(-s)},$$

les transferts  $\widetilde{M}$ ,  $\widetilde{M}^*$ ,  $\widetilde{N}$  et  $\widetilde{N}^*$  définis précédemment sont alors propres et stables et vérifient :

$$\widetilde{M}(s) \widetilde{M}^*(s) + \widetilde{N}(s) \widetilde{N}^*(s) = 1.$$

Nous obtenons alors une factorisation copremière normalisée de  $G$  (voir Définition 2.5).

Nous avons montré une méthode permettant de paramétrer les solutions d'une équation  $\mathcal{R} = 0$  suivant un système polynomial donné par (3.29). Cette première méthode nécessite des réductions modulo  $\langle g, \mathcal{B} \rangle$  et une inversion. Dans la partie suivante, nous considérons une autre méthode, plus directe, nous permettant de nous affranchir de ces opérations coûteuses.

### 3.3.2 Paramétrisation des solutions de $\mathcal{R} = 0$ par une étude directe du système d'équation polynomiales

Considérons la matrice  $\mathcal{R}$  définie par l'équation (3.11). Notons ses éléments  $\mathcal{R}_{i,j}$ . Nous notons  $X$  une solution symétrique de taille  $n \times n$  de cette équation matricielle, et par  $x_{i,j}$  ses coefficients.  $\mathcal{R}$  vérifie la proposition suivante.

**Proposition 3.9.** [RBQQ16] Soient  $(A, B, C)$  définies par (3.9). L'équation de Riccati algébrique

$$\mathcal{R} := X A + A^T X - X B B^T X + C^T C = 0$$

est équivalente au système polynomial de  $\frac{n(n+1)}{2}$  équations suivant :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n,$$

$$\mathcal{R}_{i,j} = -(x_{i,n} + a_{i-1}) (x_{j,n} + a_{j-1}) + a_{i-1} a_{j-1} + c_{i-1} c_{j-1} + x_{i,(j-1)} + x_{(i-1),j} = 0.$$

*Démonstration.* Comme  $X$  est symétrique, nous pouvons noter sa structure de la manière suivante

$$X := \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,(n-1)} & x_{1,n} \\ x_{1,2} & x_{2,2} & \dots & x_{2,(n-1)} & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1,(n-1)} & x_{2,(n-1)} & \dots & x_{(n-1),(n-1)} & x_{(n-1),n} \\ x_{1,n} & x_{2,n} & \dots & x_{(n-1),n} & x_{n,n} \end{pmatrix},$$

en d'autres termes :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad x_{j,i} = x_{i,j}.$$

En effectuant un produit matriciel, nous déduisons :

$$\begin{aligned}
 XA &= \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,(n-1)} & x_{1,n} \\ x_{1,2} & x_{2,2} & \cdots & x_{2,(n-1)} & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1,(n-1)} & x_{2,(n-1)} & \cdots & x_{(n-1),(n-1)} & x_{(n-1),n} \\ x_{1,n} & x_{2,n} & \cdots & x_{(n-1),n} & x_{n,n} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & \cdots & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix} \\
 &= \begin{pmatrix} -a_0 x_{1,n} & x_{1,1} - a_1 x_{1,n} & \cdots & x_{1,(n-1)} - a_{n-1} x_{1,n} \\ -a_0 x_{2,n} & x_{1,2} - a_1 x_{2,n} & \cdots & x_{2,(n-1)} - a_{n-1} x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_0 x_{n,n} & x_{1,n} - a_1 x_{n,n} & \cdots & x_{(n-1),n} - a_{n-1} x_{n,n} \end{pmatrix}.
 \end{aligned}$$

En adoptant les notations suivantes

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i, j \leq n, \quad \begin{cases} x_{i,0} := 0, \\ x_{0,j} := 0, \end{cases}$$

nous pouvons réécrire les éléments de  $XA$  comme suit :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad (XA)_{i,j} = x_{i,(j-1)} - a_{j-1} x_{i,n}.$$

De manière similaire, nous avons

$$\begin{aligned}
 A^T X &= \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,(n-1)} & x_{1,n} \\ x_{1,2} & x_{2,2} & \cdots & x_{2,(n-1)} & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1,(n-1)} & x_{2,(n-1)} & \cdots & x_{(n-1),(n-1)} & x_{(n-1),n} \\ x_{1,n} & x_{2,n} & \cdots & x_{(n-1),n} & x_{n,n} \end{pmatrix} \\
 &= \begin{pmatrix} -a_0 x_{1,n} & -a_0 x_{2,n} & \cdots & -a_0 x_{n,n} \\ x_{1,1} - a_1 x_{1,n} & x_{1,2} - a_1 x_{2,n} & \cdots & x_{1,n} - a_1 x_{n,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,(n-1)} - a_{n-1} x_{1,n} & x_{2,(n-1)} - a_{n-1} x_{2,n} & \cdots & x_{(n-1),n} - a_{n-1} x_{n,n} \end{pmatrix},
 \end{aligned}$$

dont nous pouvons réécrire les éléments comme suit :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad (A^T X)_{i,j} = x_{(i-1),j} - a_{i-1} x_{j,n}.$$

Ainsi, nous avons

$$\begin{aligned}
 &XA + A^T X = \\
 &\begin{pmatrix} -2a_0 x_{1,n} & x_{1,1} - a_1 x_{1,n} - a_0 x_{2,n} & x_{1,2} - a_2 x_{1,n} - a_0 x_{3,n} & \cdots & x_{1,(n-1)} - a_{n-1} x_{1,n} - a_0 x_{n,n} \\ * & 2(x_{1,2} - a_1 x_{2,n}) & x_{1,3} + x_{2,2} - a_2 x_{2,n} - a_1 x_{3,n} & \cdots & x_{2,(n-1)} - a_{n-1} x_{2,n} + x_{1,n} - a_1 x_{n,n} \\ * & * & 2(x_{2,3} - a_2 x_{3,n}) & \cdots & x_{3,(n-1)} - a_{n-1} x_{3,n} + x_{2,n} - a_2 x_{n,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & 2(x_{(n-1),n} - a_{n-1} x_{n,n}) \end{pmatrix},
 \end{aligned}$$

ou de manière équivalente :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad (XA + A^T X)_{i,j} = x_{i,(j-1)} + x_{(i-1),j} - a_{j-1} x_{i,n} - a_{i-1} x_{j,n}.$$

Par ailleurs, nous avons

$$B^T X = (0 \dots 0 \ 1) \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,(n-1)} & x_{1,n} \\ x_{1,2} & x_{2,2} & \dots & x_{2,(n-1)} & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1,(n-1)} & x_{2,(n-1)} & \dots & x_{(n-1),(n-1)} & x_{(n-1),n} \\ x_{1,n} & x_{2,n} & \dots & x_{(n-1),n} & x_{n,n} \end{pmatrix} = (x_{1,n} \dots x_{n,n}),$$

et par conséquent :

$$X B B^T X = \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{n,n} \end{pmatrix} (x_{1,n} \dots x_{n,n}) = \begin{pmatrix} x_{1,n}^2 & x_{1,n} x_{2,n} & \dots & x_{1,n} x_{n,n} \\ x_{1,n} x_{2,n} & x_{2,n}^2 & \dots & x_{2,n} x_{n,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,n} x_{n,n} & x_{2,n} x_{n,n} & \dots & x_{n,n}^2 \end{pmatrix}.$$

Les éléments de cette dernière matrice peuvent s'écrire sous la forme :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad (X B B^T X)_{i,j} = x_{i,n} x_{j,n}.$$

Enfin, nous avons

$$C^T C = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} (c_0 \ c_1 \ \dots \ c_{n-1}) = \begin{pmatrix} c_0^2 & c_0 c_1 & \dots & c_0 c_{n-1} \\ c_0 c_1 & c_1^2 & \dots & c_1 c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_0 c_{n-1} & c_1 c_{n-1} & \dots & c_{n-1}^2 \end{pmatrix}, \quad (3.33)$$

ce qui peut se traduire par :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i \leq j \leq n, \quad (C^T C)_{i,j} = c_{i-1} c_{j-1}.$$

Ainsi, les éléments de  $\mathcal{R}$  peuvent s'écrire

$$\begin{aligned} \mathcal{R}_{i,j} &= -x_{i,n} x_{j,n} - a_{j-1} x_{i,n} - a_{i-1} x_{j,n} + x_{i,(j-1)} + x_{(i-1),j} + c_{i-1} c_{j-1} \\ &= -(x_{i,n} + a_{i-1}) (x_{j,n} + a_{j-1}) + a_{i-1} a_{j-1} + c_{i-1} c_{j-1} + x_{i,(j-1)} + x_{(i-1),j} = 0, \end{aligned} \quad (3.34)$$

où  $(i, j) \in \mathbb{N}^2$  sont tels que  $1 \leq i \leq j \leq n$ . □

Examinons plus en détail le système d'équations (3.34). Pour cela, introduisons la notation suivante :

$$\begin{cases} k < 0, & b_k := 0, \\ 0 \leq k \leq n-1, & b_k := x_{(k+1),n} + a_k, \\ k = n, & b_k := 1, \\ k > n, & b_k := 0. \end{cases} \quad (3.35)$$

Adoptons aussi les notations suivantes qui nous permettent d'utiliser des indices fictifs suivants :

$$\forall k \in \mathbb{N}, \quad 1 \leq k \leq n, \quad \begin{cases} \forall i \in \mathbb{Z}, & i \leq 0, & x_{k,i} = x_{i,k} = 0, \\ \forall j \in \mathbb{N}, & j > n, & x_{j,k} = x_{k,j} = 0. \end{cases} \quad (3.36)$$

En utilisant ces notations, nous pouvons réécrire chaque élément  $\mathcal{R}_{i,j}$  (voir (3.34)) comme suit :

$$\mathcal{R}_{i,j} = -b_{i-1} b_{j-1} + a_{i-1} a_{j-1} + c_{i-1} c_{j-1} + x_{i,(j-1)} + x_{(i-1),j} = 0, \quad 1 \leq i \leq j \leq n.$$

Comme schématisé dans la Figure 3.4, nous pouvons exprimer le terme à la ligne  $i$  et la colonne  $j - 1$  par le terme à la ligne précédente  $i - 1$  et la colonne suivante  $j$  :

$$x_{i,(j-1)} = -x_{(i-1),j} + b_{i-1} b_{j-1} - (a_{i-1} a_{j-1} + c_{i-1} c_{j-1}), \quad 1 \leq i \leq j \leq n. \quad (3.37)$$

De cette remarque, nous pouvons déduire que chaque  $x_{i,j}$  peut se réécrire en fonction des  $b_k$ , c'est-à-dire en fonction de la dernière colonne de  $X$  à une translation près, comme résumé par la proposition suivante.

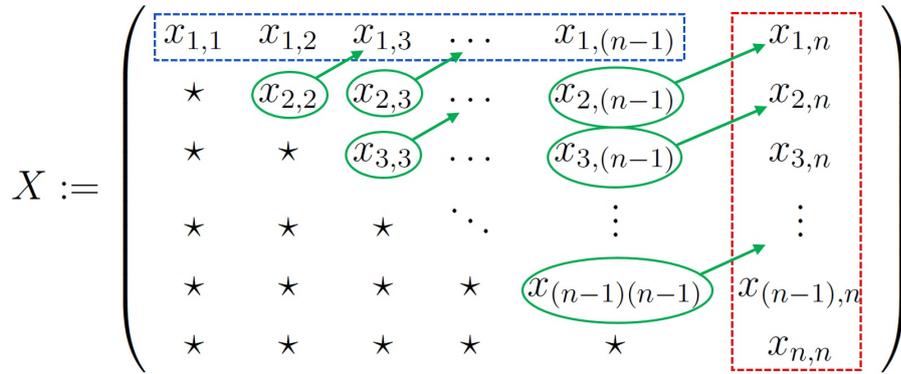


FIGURE 3.4 – Schéma de substitutions des éléments de  $X$

**Théorème 3.11.** [RBQQ16] Soit  $(i, j) \in \mathbb{N}^2$  tel que  $1 \leq i < j \leq n$ . En reprenant les notations (3.10) et (3.35), l'élément  $x_{i,j-1}$  de  $X$ , solution de  $\mathcal{R} = 0$ , peut se réécrire comme suit :

$$x_{i,(j-1)} = \sum_{k=0}^{n-1} (-1)^k (b_{i-1-k} b_{j-1+k} - (a_{i-1-k} a_{j-1+k} + c_{i-1-k} c_{j-1+k})). \quad (3.38)$$

*Démonstration.* La matrice  $X$  étant symétrique, nous pouvons nous limiter à la partie supérieure de  $X$ , c'est-à-dire ne considérer que les éléments  $x_{i,j}$  où  $i \leq j$ . Aussi, de part la définition de  $b_k$  donnée par (3.35), nous tirons trivialement :

$$\forall k \in \llbracket 1; n \rrbracket, \quad x_{k,n} = b_{k-1} - a_{k-1}. \quad (3.39)$$

Il ne reste alors plus qu'à exprimer les  $x_{i,j}$  pour  $1 \leq i \leq j < n$ . De (3.37), en opérant  $m$  substitutions successives, nous obtenons :

$$\forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i < j \leq n, \\ x_{i,(j-1)} = -(-1)^m x_{(i-1-m),(j+m)} + \sum_{k=0}^m (-1)^k (b_{i-1-k} b_{j-1+k} - (a_{i-1-k} a_{j-1+k} + c_{i-1-k} c_{j-1+k})). \quad (3.40)$$

Le schéma de substitution des  $x_{i,j}$  s'arrête quand nous atteignons soit la première ligne (c'est-à-dire  $m = i - 2$ ), soit la dernière colonne de  $X$  (c'est-à-dire  $m = n - j$ ). En effet, d'après (3.39) la dernière colonne de  $X$  s'exprime directement en fonction des  $b_k$ . En particulierisant

(3.37) par  $i = 1$  et en utilisant que  $x_{0,j} = 0$ , nous obtenons alors que tous les éléments de la première ligne s'écrivent en fonction des  $b_k$  :

$$x_{1,(j-1)} = b_0 b_{j-1} - (a_0 a_{j-1} + c_0 c_{j-1}), \quad j = 2, \dots, n. \quad (3.41)$$

Afin d'alléger l'écriture, nous considérons les notations (3.35) et (3.36) qui prennent en compte des indices négatifs. Si nous effectuons  $n - 1$  substitutions, c'est-à-dire en particulier  $m = n - 1$  dans (3.40), nous obtenons alors :

$$\begin{aligned} \forall (i, j) \in \mathbb{N}^2, \quad 1 \leq i < j \leq n, \\ x_{i,(j-1)} &= (-1)^n x_{(i-n),(j+n-1)} + \sum_{k=0}^{n-1} (-1)^k (b_{i-1-k} b_{j-1+k} - (a_{i-1-k} a_{j-1+k} + c_{i-1-k} c_{j-1+k})) \\ &= \sum_{k=0}^{n-1} (-1)^k (b_{i-1-k} b_{j-1+k} - (a_{i-1-k} a_{j-1+k} + c_{i-1-k} c_{j-1+k})). \end{aligned}$$

Nous obtenons alors l'expression (3.38), ce qui conclut la preuve du théorème.  $\square$

Notons que (3.38) est obtenu en considérant les équations  $\mathcal{R}_{i,j} = 0$  pour  $1 \leq i < j \leq n$ . Il reste  $n$  équations non exploitées, c'est-à-dire les éléments diagonaux  $\mathcal{R}_{i,i} = 0$  pour  $1 \leq i \leq n$ . Notons aussi que nous ne savons pas encore quelles équations doivent vérifier les  $b_k$ . La proposition suivante permet de répondre à cette dernière question en utilisant les  $n$  équations restantes et le Théorème 3.11.

**Proposition 3.10.** [RBQQ16] Notons par  $d_{2k}$  les combinaisons de paramètres  $a_k, c_k$  définis par (3.10), c'est-à-dire :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad d_{2k} := a_k^2 + c_k^2 + 2 \sum_{m=1}^n (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}). \quad (3.42)$$

Soient les polynômes suivants :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad \mathcal{B}_k := b_k^2 + 2 \sum_{m=1}^n (-1)^m b_{k-m} b_{k+m} - d_{2k}. \quad (3.43)$$

Notons par  $\mathcal{B}$  l'ensemble des polynômes :

$$\mathcal{B} := \{\mathcal{B}_k\}_{k=0, \dots, n-1}. \quad (3.44)$$

Les  $n$  variables  $b_k$  définies en (3.35) sont solutions du système polynomial  $\mathcal{E}$  défini par :

$$\mathcal{E} := \{\mathcal{B}_k = 0\}_{k=0, \dots, n-1}. \quad (3.45)$$

*Démonstration.* En particulierisant (3.37) en  $j = i$  et en utilisant que  $x_{(i-1),i} = x_{i,(i-1)}$ , nous obtenons :

$$x_{(i-1),i} = x_{i,(i-1)} = \frac{1}{2} (b_{i-1}^2 - (a_{i-1}^2 + c_{i-1}^2)), \quad i = 1, \dots, n. \quad (3.46)$$

Tout d'abord, notons que pour  $i = 1$ , (3.46) devient

$$\mathcal{B}_0 := b_0^2 - (a_0^2 + c_0^2) = 0, \quad (3.47)$$

c'est-à-dire une équation dépendant uniquement des  $a_i, b_j$  et  $c_k$ .

Pour  $i \neq 1$ , nous pouvons utiliser (3.38) et (3.46) afin d'en déduire des équations sur les  $b_k$ . (3.38) n'est utilisable qu'à la condition que  $i < j$ . En substituant  $i = l - 1$  et  $j = l + 1$  dans (3.38), nous obtenons :

$$l = 2, \dots, n - 1, \quad x_{(l-1),l} = \sum_{k=0}^{n-1} (-1)^k (b_{l-2-k} b_{l+k} - (a_{l-2-k} a_{l+k} + c_{l-2-k} c_{l+k})). \quad (3.48)$$

En combinant (3.46) et (3.48), nous obtenons  $n - 2$  nouvelles équations dépendant uniquement de  $a_i$ ,  $b_j$  et  $c_k$  :

$$l = 2, \dots, n - 1, \quad \frac{1}{2} (b_{l-1}^2 - (a_{l-1}^2 + c_{l-1}^2)) = \sum_{k=0}^{n-1} (-1)^k (b_{l-2-k} b_{l+k} - (a_{l-2-k} a_{l+k} + c_{l-2-k} c_{l+k})). \quad (3.49)$$

(3.49) peut alors se réécrire :

$$l = 2, \dots, n - 1, \quad \mathcal{B}_{l-1} = 0,$$

où

$$\mathcal{B}_{l-1} := b_{l-1}^2 - 2 \sum_{k=0}^{n-1} (-1)^k b_{l-2-k} b_{l+k} - \left( a_{l-1}^2 + c_{l-1}^2 - 2 \sum_{k=0}^{n-1} (-1)^k (a_{l-2-k} a_{l+k} + c_{l-2-k} c_{l+k}) \right) = 0.$$

Substituons  $l - 1$  par  $l$  dans  $\mathcal{B}_{l-1}$  :

$$l = 1, \dots, n - 2, \quad \mathcal{B}_l := b_l^2 - 2 \sum_{k=0}^{n-1} (-1)^k b_{l-1-k} b_{l+1+k} - \left( a_l^2 + c_l^2 - 2 \sum_{k=0}^{n-1} (-1)^k (a_{l-1-k} a_{l+1+k} + c_{l-1-k} c_{l+1+k}) \right).$$

En substituant  $k + 1$  par  $k$ , nous déduisons l'écriture suivante :

$$\forall l \in \mathbb{N}, \quad 1 \leq l \leq n - 2, \quad \mathcal{B}_l := b_l^2 + 2 \sum_{k=1}^n (-1)^k b_{l-k} b_{l+k} - \left( a_l^2 + c_l^2 + 2 \sum_{k=1}^n (-1)^k (a_{l-k} a_{l+k} + c_{l-k} c_{l+k}) \right). \quad (3.50)$$

Enfin, pour  $i = n$ , (3.46) devient

$$x_{(n-1),n} = \frac{1}{2} (b_{n-1}^2 - (a_{n-1}^2 + c_{n-1}^2)),$$

soit, en utilisant  $x_{(n-1),n} = b_{n-2} - a_{n-2}$  et que  $b_n = 1$  et  $a_n = 1$  :

$$\mathcal{B}_{n-1} := b_{n-1}^2 - 2 b_n b_{n-2} - (a_{n-1}^2 + c_{n-1}^2 - 2 a_n a_{n-2}). \quad (3.51)$$

Enfin, remarquons qu'en adoptant les notations (3.35), l'écriture (3.50) permet d'exprimer (3.47) en  $l = 0$  et (3.51) en  $l = n - 1$ . Nous choisissons alors de conserver l'unique écriture (3.50). En utilisant la notation

$$d_{2l} = a_l^2 + c_l^2 + 2 \sum_{k=1}^n (-1)^k (a_{l-k} a_{l+k} + c_{l-k} c_{l+k}),$$

nous avons

$$\mathcal{B}_l := b_l^2 + 2 \sum_{k=1}^n (-1)^k b_{l-k} b_{l+k} - d_{2l},$$

ce qui conclut la preuve. □

**Remarque 3.15.** Les coefficients  $a$  et  $c$  n'apparaissent que dans les constantes  $d_{2k}$  des polynômes  $\mathcal{B}_k$ . De plus, les coefficients des termes quadratiques en les  $b_k$  des polynômes  $\mathcal{B}_k$  sont 1 ou  $2(-1)^m$ . En particulier, le degré des  $\mathcal{B}_k$  est 2 quel que soient les valeurs des paramètres  $a$  et  $c$ .

**Remarque 3.16.** Pour tout système SISO d'ordre  $n \in \mathbb{N}$ , l'association du Théorème 3.11 et de la Proposition 3.10 montre que nous pouvons réduire la résolution de l'équation de Riccati algébrique  $\mathcal{R} = 0$  à celle d'un système polynomial de taille  $n$ . Cette réduction est immédiate, par application directe des formules explicites (3.38) et (3.43).

**Exemple 3.11.** Reprenons le système de l'Exemple 3.10 de fonction de transfert :

$$G := \frac{c_0}{s^2 + a_1 s + a_0}, \quad c_0 := \frac{1}{m}, \quad a_1 := \frac{\xi}{m}, \quad a_0 := \frac{k}{m}.$$

D'après (3.9), l'ERA associée à ce système  $G$  s'écrit de la manière suivante :

$$\mathcal{R} = \begin{pmatrix} -2a_0 x_{1,2} + c_0^2 - x_{1,2}^2 & -a_0 x_{2,2} - a_1 x_{1,2} - x_{1,2} x_{2,2} + x_{1,1} \\ -a_0 x_{2,2} - a_1 x_{1,2} - x_{1,2} x_{2,2} + x_{1,1} & -2a_1 x_{2,2} - x_{2,2}^2 + 2x_{1,2} \end{pmatrix} = \mathbb{O}_{2 \times 2}.$$

En utilisant les notations (3.35), nous obtenons :

$$\begin{cases} x_{1,2} = b_0 - a_0, \\ x_{2,2} = b_1 - a_1. \end{cases}$$

L'équation  $\mathcal{R} = 0$  s'écrit alors en fonction de  $a_k, b_k, c_k$  comme suit :

$$\mathcal{R} = \begin{pmatrix} a_0^2 - b_0^2 + c_0^2 & a_1 a_0 - b_0 b_1 + x_{1,1} \\ a_1 a_0 - b_0 b_1 + x_{1,1} & a_1^2 - b_1^2 - 2a_0 + 2b_0 \end{pmatrix} = \mathbb{O}_{2 \times 2}.$$

Il est alors facile d'extraire  $x_{1,1}$  de  $\mathcal{R}_{1,2} = 0$ ,

$$x_{1,1} = -a_0 a_1 + b_0 b_1,$$

et donc d'obtenir la forme suivante de  $X$  solution de  $\mathcal{R} = 0$  :

$$X = \begin{pmatrix} -a_0 a_1 + b_0 b_1 & b_0 - a_0 \\ b_0 - a_0 & b_1 - a_1 \end{pmatrix}.$$

Notons que nous retrouvons ce résultat en appliquant directement le Théorème 3.11. De plus, notons que les termes diagonaux de  $\mathcal{R}$  dépendent uniquement de  $b_0$  et  $b_1$ . De ces éléments diagonaux, nous obtenons directement le système suivant :

$$\begin{cases} b_0^2 = a_0^2 + c_0^2, \\ b_1^2 + 2b_0 = a_1^2 + c_1^2 + 2a_0. \end{cases}$$

Nous retrouvons alors le système  $\mathcal{B}$  défini à la Proposition 3.10.

**Exemple 3.12.** Afin d'expliciter le schéma de substitution intervenant dans l'ERA, considérons un modèle de deux masses  $m_1$  et  $m_2$  reliées par un ressort de raideur  $k$  sans d'amortissement. Ce système est un exemple standard en théorie du contrôle [WB92, Vin00, ACA<sup>+</sup>99]. Celui-ci est représenté en Figure 3.5.

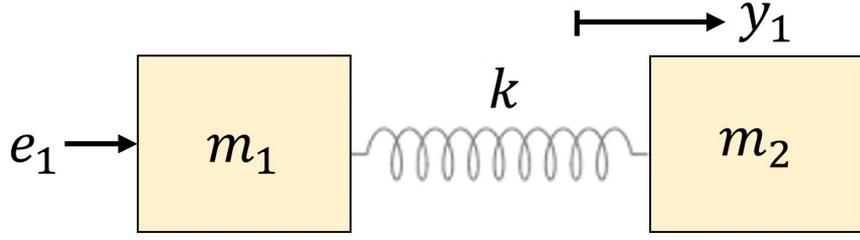


FIGURE 3.5 – Modèle deux-masses-ressort

Avec les notations de la Figure 3.5, nous étudions le déplacement de la deuxième masse, notée  $y_1$ , lorsque la première est soumise à une force  $e_1$  l'entraînant en translation. Nous noterons  $m = m_1 + m_2$  et nous considérons alors le système  $G$  entre l'entrée  $e_1$  et la sortie  $y_1$ , de fonction de transfert :

$$G = \frac{y_1}{e_1} = \frac{c_0}{s^2 (s^2 + a_2)}, \quad a_2 := \frac{m k}{m_1 m_2}, \quad c_0 := \frac{a_2}{m} > 0.$$

D'après (3.9), l'ERA associée à ce système  $G$  s'écrit de la manière suivante :

$$\mathcal{R} = \begin{pmatrix} c_0^2 - x_{1,4}^2 & -x_{1,4} x_{2,4} + x_{1,1} & -a_2 x_{1,4} - x_{1,4} x_{3,4} + x_{1,2} & -x_{1,4} x_{4,4} + x_{1,3} \\ \bullet & -x_{2,4}^2 + 2 x_{1,2} & -a_2 x_{2,4} - x_{2,4} x_{3,4} + x_{1,3} + x_{2,2} & -x_{2,4} x_{4,4} + x_{1,4} + x_{2,3} \\ \bullet & \bullet & -2 a_2 x_{3,4} - x_{3,4}^2 + 2 x_{2,3} & -a_2 x_{4,4} - x_{3,4} x_{4,4} + x_{2,4} + x_{3,3} \\ \bullet & \bullet & \bullet & -x_{4,4}^2 + 2 x_{3,4} \end{pmatrix} = \mathbb{O}_{4 \times 4}.$$

En utilisant les notations (3.35), nous obtenons :

$$\begin{cases} x_{1,4} = b_0, \\ x_{2,4} = b_1, \\ x_{3,4} = b_2 - a_2, \\ x_{4,4} = b_3. \end{cases}$$

L'équation  $\mathcal{R} = 0$  s'écrit alors en fonction de  $a_k, b_k, c_k$  comme suit :

$$\mathcal{R} := \begin{pmatrix} -b_0^2 + c_0^2 & -b_0 b_1 + x_{1,1} & -b_0 b_2 + x_{1,2} & -b_0 b_3 + x_{1,3} \\ \bullet & -b_1^2 + 2 x_{1,2} & -b_1 b_2 + x_{1,3} + x_{2,2} & -b_1 b_3 + b_0 + x_{2,3} \\ \bullet & \bullet & a_2^2 - b_2^2 + 2 x_{2,3} & -b_3 b_2 + b_1 + x_{3,3} \\ \bullet & \bullet & \bullet & -b_3^2 - 2 a_2 + 2 b_2 \end{pmatrix} = \mathbb{O}_{4 \times 4}.$$

Nous remarquons alors que :

- Nous pouvons tirer  $x_{1,1}$  de  $\mathcal{R}_{1,2} = 0$  et  $x_{1,2}$  de  $\mathcal{R}_{1,3} = 0$  :

$$\begin{cases} x_{1,1} = b_0 b_1, \\ x_{1,2} = b_0 b_2. \end{cases}$$

- Nous pouvons tirer  $x_{1,3}$  de  $\mathcal{R}_{2,3} = 0$  en fonction de  $x_{2,2}$  :

$$x_{1,3} = b_1 b_2 - x_{2,2}.$$

- En injectant l’expression précédente de  $x_{1,3}$  dans  $\mathcal{R}_{1,4}$ , nous tirons  $x_{2,2}$  :

$$x_{2,2} = -b_0 b_3 + b_1 b_2.$$

De plus, nous pouvons tirer  $x_{2,3}$  de  $\mathcal{R}_{2,4} = 0$  et  $x_{3,3}$  de  $\mathcal{R}_{3,4} = 0$  :

$$\begin{cases} x_{2,3} = b_1 b_3 - b_0, \\ x_{3,3} = b_2 b_3 - b_1. \end{cases}$$

- Il reste alors 4 équations à exploiter correspondant à la diagonale de  $\mathcal{R}$ . Ces équations font intervenir  $x_{1,2}$  et  $x_{2,3}$  dont nous disposons des expressions dépendant des  $a_k$ ,  $b_k$  et  $c_k$ . Ainsi, nous retrouvons le système  $\mathcal{B}$  à partir de la diagonale de  $\mathcal{R}$ , c’est-à-dire

$$\begin{cases} \mathcal{B}_0 := b_0^2 - d_0 = 0, \\ \mathcal{B}_1 := b_1^2 - 2 b_0 b_2 = 0, \\ \mathcal{B}_2 := b_2^2 - 2 b_1 b_3 + 2 b_0 - d_4 = 0, \\ \mathcal{B}_3 := b_3^2 - 2 b_2 - d_6 = 0, \end{cases}$$

où :

$$\begin{cases} d_0 := c_0^2, \\ d_4 := a_2^2, \\ d_6 := -2 a_2. \end{cases}$$

### 3.3.3 Équivalence des deux approches

D’un côté, [KYAH09] propose une méthode de résolution symbolique en passant par une factorisation spectrale du polynôme caractéristique de l’Hamiltonien  $\mathcal{H}$ . De l’autre, nous proposons une méthode plus directe en étudiant le système d’équations polynomiales qui découle de l’ERA. Nous montrons que ces deux méthodes sont équivalentes en démontrant qu’elles impliquent la résolution du même système d’équations polynomiales.

**Théorème 3.12.** Soient  $k \in \mathbb{Z}$ ,  $\beta_k$  défini par la factorisation spectrale de  $f$  donné à la Définition 3.21 et  $b_k$  défini par (3.35). Alors, nous avons  $\beta_k = b_k$ . De plus, nous avons  $\mathcal{B} = \mathcal{B}$ , où  $\mathcal{B}$  est défini en (3.29) et  $\mathcal{B}$  est défini en (3.44).

*Démonstration.* Ce théorème découle directement de la comparaison des systèmes (3.29) et (3.43).  $\square$

Par la suite, cette équivalence sera utilisée, notamment pour caractériser la solution définie positive de l’équation de Riccati algébrique.

Le système polynomial obtenu  $\mathcal{E}$  a des propriétés particulières que nous allons exploiter.

## 3.4 Propriétés du système polynomial $\mathcal{E}$ issu de $\mathcal{R} = 0$

La section précédente montre que chaque composante  $x_{i,j}$  d’une solution  $X$  de  $\mathcal{R} = 0$  peut s’écrire en fonction des  $b_k$  définis par (3.35), et ces mêmes  $b_k$  satisfont un système d’équations polynomiales  $\mathcal{B}$  défini par

$$\mathcal{E} := \{\mathcal{B}_k = 0 \mid 0 \leq k \leq n - 1\},$$

où :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad d_{2k} := a_k^2 + c_k^2 + 2 \sum_{m=1}^n (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}),$$

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad \mathcal{B}_k := b_k^2 + 2 \sum_{m=1}^n (-1)^m b_{k-m} b_{k+m} - d_{2k}.$$

De plus, notons :

$$\mathcal{B} := (\mathcal{B}_0, \dots, \mathcal{B}_{n-1}).$$

Nous noterons aussi par  $\mathbb{K} := \mathbb{Q}(d_0, \dots, d_{2n})$  le corps des fractions rationnelles en les coefficients  $d_{2k}$ , et  $R := \mathbb{K}[b_0, \dots, b_{n-1}]$  l'anneau des polynômes en les variables  $b_0, \dots, b_{n-1}$  à coefficients dans  $\mathbb{K}$ . De plus, notons  $\langle \mathcal{B} \rangle$  l'idéal engendré par  $\mathcal{B}$  et la  $\mathbb{K}$ -algèbre  $\mathcal{A} := R/\langle \mathcal{B} \rangle$ .

**Théorème 3.13.** [RBQQ16] La  $\mathbb{K}$ -algèbre  $\mathcal{A} := R/\langle \mathcal{B} \rangle$  est *zéro-dimensionnelle*, c'est-à-dire forme un  $\mathbb{K}$ -espace vectoriel de dimension finie. Plus précisément, sa dimension est  $2^n$ . De plus,  $\mathcal{B}$  est une base de Gröbner réduite de l'idéal  $\langle \mathcal{B} \rangle$  pour l'ordre monomial du degré lexicographique inverse (DRL)  $b_{n-1} \succ \dots \succ b_0$ .

*Démonstration.* Une application directe du *critère de Buchberger* (voir Théorème 3.6 de la Section 3.1.7 ou [CLO15, Th. 6, §6, Chap. 2]) montre que  $\mathcal{B}$  est une base de Gröbner de  $\langle \mathcal{B} \rangle$  suivant l'ordre du degré lexicographique inverse (DRL)  $b_{n-1} \succ \dots \succ b_0$ . De plus, par définition (voir Définition 3.16 de la Section 3.1.7 ou [CLO15, Def. 4, §7, Chap. 2]), cette base de Gröbner est réduite.

Comme le coefficient de tête de chaque  $\mathcal{B}_k$  vaut 1 (ne dépend pas de  $d_{2k}$ ), le monôme de tête de chaque polynôme de la base de Gröbner  $\mathcal{B}$  est de la forme  $b_k^2$ ,  $\mathcal{A}$  est un  $\mathbb{K}$ -espace vectoriel de dimension finie [CLO15, Th. 4, §6, Chap. 5]. En utilisant la notation  $b^\alpha := \prod_{i=0}^{n-1} b_i^{\alpha_i}$  où  $\alpha := (\alpha_0, \dots, \alpha_{n-1}) \in \mathbb{N}^n$ , une base de  $\mathcal{A}$  est définie par  $\{b^\alpha\}_{\alpha \in \llbracket 0,1 \rrbracket^n}$ . Clairement, la dimension sur  $\mathbb{K}$  de  $\mathcal{A}$  est  $2^n$ , ce qui conclut la preuve.  $\square$

**Remarque 3.17.** A paramètres  $a$  et  $c$  sont fixés, le fait que  $\mathcal{A}$  soit zéro-dimensionnel implique que le système  $\mathcal{E}$  a un nombre fini de solutions sur  $\mathbb{C}$  comptées avec multiplicité (voir [CLO15]).

**Corollaire 3.1.** Le système polynomial  $\{\mathcal{R}_{i,j} = 0\}$  associé à l'équation de Riccati algébrique d'un système SISO est un  $\mathbb{K}$ -espace vectoriel de dimension finie, de dimension  $2^n$ , et de base  $\{b^\alpha\}_{\alpha \in \llbracket 0,1 \rrbracket^n}$ .

*Démonstration.* La preuve découle immédiatement des Théorèmes 3.11 et 3.13.  $\square$

Réduire l'équation de Riccati algébrique au système polynomial  $\mathcal{E}$  donne directement une base de Gröbner réduite de l'idéal  $\langle \mathcal{B} \rangle$ . Étant donné cette base de Gröbner réduite, nous pouvons alors déjà utiliser l'algorithme de division pour simplifier des polynômes par rapport à  $\mathcal{E}$  avec garantie d'unicité du résultat. Cette propriété est extrêmement utile notamment pour calculer une paramétrisation de  $V(\langle \mathcal{B} \rangle)$ . Pour des paramètres  $a$  et  $c$  non fixés, cette dernière variété est incluse dans une clôture algébrique difficilement manipulable. Ainsi, nous proposons dans un premier temps de travailler à paramètres fixés, puis d'étendre les résultats à paramètres non fixés.

## 3.5 Étude de $\mathcal{E}$ à paramètres fixés

### 3.5.1 Exemple d'introduction

À paramètres  $a$  et  $c$  fixés, le système  $\mathcal{E}$  possède un nombre fini de solutions, et ces solutions peuvent se représenter par un  $n$ -uplet  $(b_0, \dots, b_{n-1})$ , soit un point de  $\mathbb{C}^n$ . Ces points représentent la variété algébrique  $V(\langle \mathcal{B} \rangle)$ . Illustrons ceci par un exemple.

**Exemple 3.13.** Considérons  $n = 2$ ,  $d_0 = \frac{1}{4}$  et  $d_2 = 2$ . Nous avons :

$$\mathcal{E} = \begin{cases} b_0^2 = \frac{1}{4}, \\ b_1^2 - 2b_0 = 2, \end{cases} \Leftrightarrow \begin{cases} b_0 = \pm \frac{1}{2}, \\ b_1^2 = 2 + 2b_0. \end{cases}$$

En d'autres termes, il y a quatre couples  $(b_0, b_1)$  solutions dans  $\mathbb{C}$  de  $\mathcal{E}$  :  $(-\frac{1}{2}, 1)$ ,  $(-\frac{1}{2}, -1)$ ,  $(\frac{1}{2}, \sqrt{3})$ ,  $(\frac{1}{2}, -\sqrt{3})$ . Nous pouvons alors représenter ces solutions dans le plan  $(b_0, b_1)$ .

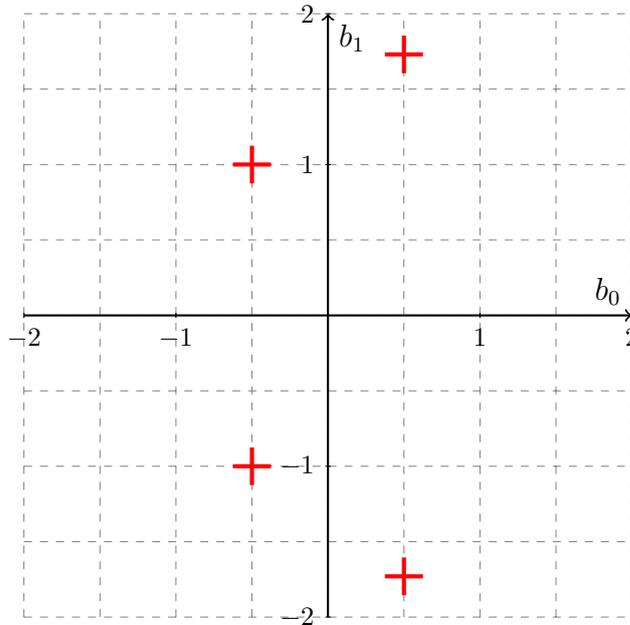


FIGURE 3.6 – Représentation de  $V(\mathcal{B})$  dans le cas  $n = 2$ ,  $d_0 = 1/4$ ,  $d_2 = 2$

Dans cet exemple, le nombre de solutions distinctes est  $2^2 = 4$ , soit exactement la dimension de  $\mathcal{A}$ . Autrement dit, chaque zéro de  $\mathcal{B}$  est de multiplicité 1 ou encore, l'idéal  $\langle \mathcal{B} \rangle$  est radical [CLO15, Chap. 4, §2]. De plus, nous pouvons voir que la coordonnée  $b_1$  de chaque solution permet d'identifier celle-ci de façon unique (voir la Figure 3.6). Nous dirons alors que  $b_1$  est *séparant*.

Supposons alors, dans cette section,  $a \in \mathbb{Q}^n$  et  $c \in \mathbb{Q}^n$ . Nous avons alors  $d_{2k} \in \mathbb{Q}$ ,  $k = 0, \dots, n$ .  $\mathbb{K}$  défini à la section précédente devient alors le corps  $\mathbb{Q}$ , et  $R$  devient  $\mathbb{Q}[b_0, \dots, b_{n-1}]$ . Plus généralement, dans ce contexte, nous pouvons définir un élément séparant, ce que nous détaillons dans la section suivante.

### 3.5.2 Éléments séparant $V(\langle \mathcal{B} \rangle)$

Nous définissons un élément séparant  $V(\langle \mathcal{B} \rangle)$  comme suit.

**Définition 3.22.** Un polynôme  $t$  dans  $\mathbb{Q}[b_0, \dots, b_{n-1}]$  est dit *séparant pour  $\langle \mathcal{B} \rangle$*  si et seulement si :

$$\forall \nu, \mu \in V(\langle \mathcal{B} \rangle), \quad \nu \neq \mu \Rightarrow t(\nu) \neq t(\mu).$$

Un tel polynôme  $t := \sum_{i=0}^{n-1} u_i b_i$ ,  $u_i \in \mathbb{Q}$ , de degré 1 est appelé *forme séparante*.

**Remarque 3.18.** La forme  $t$  est donc séparante pour les zéros de  $\mathcal{B}$  si l'application suivante est injective :

$$\begin{aligned} V(\langle \mathcal{B} \rangle) &\rightarrow \mathbb{C} \\ (b_0, \dots, b_{n-1}) &\mapsto t = \sum_{i=0}^{n-1} u_i b_i. \end{aligned}$$

Trouver une forme séparante  $\langle \mathcal{B} \rangle$  est la première étape à réaliser pour trouver une paramétrisation de  $V(\langle \mathcal{B} \rangle)$ . En effet, si une forme séparante  $t$  de  $V(\langle \mathcal{B} \rangle)$  est donnée, on peut alors exprimer chaque coordonnée de  $V(\langle \mathcal{B} \rangle)$  comme une fraction rationnelle en  $t$ , où  $t$  vérifie une équation polynomiale à coefficients dans  $\mathbb{Q}$ . La proposition suivante, bien qu'intuitive, est centrale dans notre démarche.

**Théorème 3.14.** [Rou99] Tout idéal zéro-dimensionnel sur  $\mathbb{C}$  admet au moins une forme séparante.

**Remarque 3.19.** Le Théorème 3.14 se comprend aisément d'un point de vue géométrique. Trouver une forme séparante revient à trouver une direction de projection. Afin de mieux expliquer ce point, reprenons l'Exemple 3.13 et considérons la Figure 3.7. Nous cherchons une forme séparante de la forme  $t = u_0 b_0 + u_1 b_1$  avec  $u_0, u_1 \in \mathbb{Q}$ . La direction de projection associée à cette forme séparante est la droite  $u_0 b_0 + u_1 b_1 = 0$ . Si la direction de projection de  $t$  n'est pas colinéaire à l'une des droites passant par deux points de  $V(\langle \mathcal{B} \rangle)$ , alors  $t$  est séparante. Autrement dit, si  $(b_0^{(i)}, b_1^{(i)})$  et  $(b_0^{(j)}, b_1^{(j)})$  sont deux points de  $V(\langle \mathcal{B} \rangle)$  tels que

$$u_0 b_0^{(i)} + u_1 b_1^{(i)} = u_0 b_0^{(j)} + u_1 b_1^{(j)},$$

alors  $(b_0^{(i)}, b_1^{(i)})$  et  $(b_0^{(j)}, b_1^{(j)})$  sont sur une même droite et  $t = u_0 b_0 + u_1 b_1$  n'est pas séparante. Par exemple, la forme linéaire  $t = b_1 : (b_0, b_1) \mapsto b_1$  correspond à la projection sur l'axe  $b_1$ . Sa direction de projection est l'axe  $b_0$ , qui n'est pas colinéaire à aucune des droites jaunes représentées sur la Figure 3.7. La forme linéaire  $t = b_1 : (b_0, b_1) \mapsto b_1$  est donc séparante. Un autre exemple est la forme linéaire  $t = b_0 : (b_0, b_1) \mapsto b_0$ , correspondant à la projection sur l'axe  $b_0$ . Sa direction de projection est l'axe  $b_1$ , qui est colinéaire à deux droites jaunes représentées sur la Figure 3.7. Autrement dit, deux points ont la même abscisse (par exemple, les points  $(-1/2, -1)$  et  $(-1/2, 1)$ ). La forme linéaire  $t = b_0 : (b_0, b_1) \mapsto b_0$  n'est donc pas séparante. Un dernier exemple est la forme linéaire  $t : (b_0, b_1) \mapsto b_0 - 2b_1$ . Sa direction de projection est la droite  $b_1 = \frac{1}{2} b_0$ , représentée en bleue sur la Figure 3.7. Cette droite n'est parallèle à aucune des droites jaunes. La forme linéaire  $t : (b_0, b_1) \mapsto b_0 - 2b_1$  est donc séparante. Plus généralement, nous pourrions prendre *presque* toutes les droites passant par l'origine comme direction de projection puisque la seule restriction pour obtenir une forme séparante est que la direction de projection ne soit pas parallèle à l'une des droites jaunes.

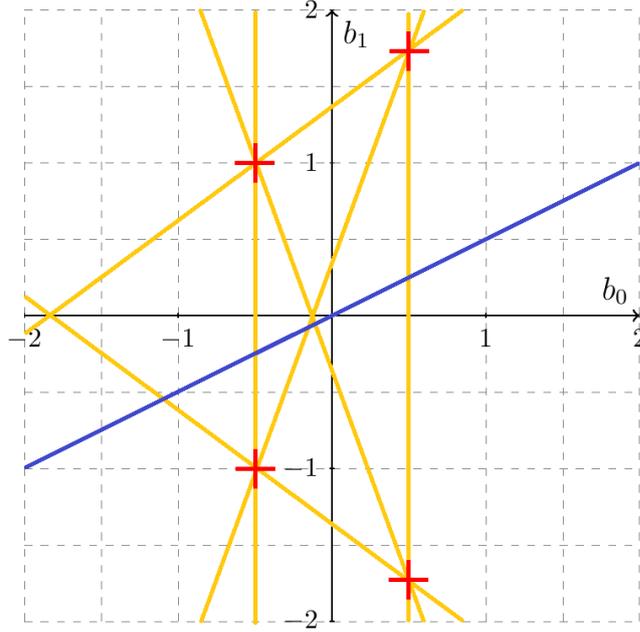


FIGURE 3.7 – Représentation de  $V(\mathcal{B})$  dans le cas  $n = 2$ ,  $d_0 = 1/4$ ,  $d_2 = 2$  (croix rouges), des directions de projection des formes linéaires non séparant  $V(\mathcal{B})$  (droites jaunes) et de la direction de projection de la forme séparante  $t : (b_0, b_1) \mapsto b_0 - 2b_1$  (droite bleue)

L'idéal  $\langle \mathcal{B} \rangle$  étant zéro-dimensionnel, il admet au moins une forme séparante. Il ne reste plus qu'à en identifier une. Quelques exemples tendent à montrer que  $b_{n-(2k-1)}$  est séparant. Il se trouve que ceci est vrai en dehors d'un *fermé de Zariski*, c'est-à-dire en dehors d'une certaine variété algébrique. Autrement dit, pour *presque toutes les valeurs de paramètres*, la variable  $b_{n-(2k-1)}$  est séparante. En accord avec la géométrie algébrique, c'est ce que nous appelons *cas générique*.

Établissons maintenant plus précisément l'affirmation précédente.

**Théorème 3.15.** Soient  $a \in \mathbb{Q}^n$  et  $c \in \mathbb{Q}^n$ . Soient  $\lambda_i \in \mathbb{C}_-$ ,  $i = 1, \dots, n$  les racines de  $f = (-1)^n \sum_{k=0}^n (-1)^k d_{2k} \lambda^{2k}$  (voir (3.13) et (3.14)) à partie réelle strictement négative, où

$$d_{2k} := a_k^2 + c_k^2 + 2 \sum_{m=1}^k (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}), \quad k = 1, \dots, n.$$

Nous avons les résultats suivants.

- Supposons  $n = 1$ . Alors  $b_0$  est séparant si et seulement si  $d_0 = a_0^2 + c_0^2 \neq 0$ , c'est-à-dire si et seulement si  $a_0 \neq 0$  ou  $c_0 \neq 0$ .
- Supposons  $n > 1$ . Alors  $b_0$  n'est pas séparant. De plus, pour  $k = 1, \dots, n/2$ ,  $b_{n-2k}$  n'est pas séparant et  $b_{n-(2k-1)}$  est séparant *génériquement*, c'est-à-dire si et seulement si  $D_k \neq 0$ , où  $D_k \in \mathbb{Q}[d_0, \dots, d_{2(n-1)}]$  est un polynôme constructible à partir des coefficients de  $f$ . Le lieu d'annulation de  $D_k$  sera appelé *cas singulier*.

*Démonstration.* Rappelons que l'on peut retrouver le système  $\mathcal{B}$  à partir d'une factorisation spectrale de l'Hamiltonien  $\mathcal{H}$  associé à l'équation de Riccati algébrique

$$f(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i)(\lambda + \lambda_i) = (-1)^n g(\lambda) g(-\lambda), \quad g(\lambda) = \sum_{i=0}^n b_i \lambda^i = \prod_{i=1}^n (\lambda - \epsilon_i \lambda_i),$$

où  $\lambda_i \in \mathbb{C}_-$  pour  $i = 1, \dots, n$  et  $\epsilon_i \in \{-1, 1\}$ . De cette factorisation spectrale et des sommes de Newton, on en déduit que :

$$\left\{ \begin{array}{l} b_{n-1} = -\sum_{i=1}^n \epsilon_i \lambda_i, \\ b_{n-2} = \sum_{1 \leq i < j \leq n} \epsilon_i \lambda_i \epsilon_j \lambda_j, \\ \vdots \\ b_{n-k} = (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \epsilon_{i_j} \lambda_{i_j}, \\ \vdots \\ b_0 = \prod_{i=1}^n \epsilon_i \lambda_i. \end{array} \right. \quad (3.52)$$

• Supposons  $n = 1$ . Alors, nous avons  $b_0 = \epsilon_1 \lambda_1$ . Ainsi,  $b_0$  prend des valeurs distinctes suivant les valeurs de  $\epsilon_1$ , sauf si  $\lambda_1 = 0$ . Ceci prouve la séparabilité de  $b_0$  dans le cas générique, la non-généricité étant caractérisée par la condition  $a_0^2 + c_0^2 = b_0^2 = 0$  qui définit un point dans l'espace réel des paramètres  $(a_0, c_0)$ .

• Supposons  $n > 1$ . Soient deux jeux de  $n$  racines de  $f$ , notées par  $\{\epsilon_1 \lambda_1, \dots, \epsilon_n \lambda_n\}$  et  $\{\epsilon'_1 \lambda_1, \dots, \epsilon'_n \lambda_n\}$ . D'après (3.52), montrer que  $b_{n-k}$  est séparant pour  $V(\langle \mathcal{B} \rangle)$  est équivalent à dire que :

$$(-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \epsilon_{i_j} \lambda_{i_j} = (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \epsilon'_{i_j} \lambda_{i_j} \quad \Rightarrow \quad \forall i \in \llbracket 1, n \rrbracket, \quad \epsilon_i = \epsilon'_i. \quad (3.53)$$

En factorisant la partie gauche de l'implication (3.53), nous obtenons :

$$\begin{aligned} (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \epsilon_{i_j} \lambda_{i_j} &= (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \epsilon'_{i_j} \lambda_{i_j} \\ &\Leftrightarrow \sum_{1 \leq i_1 < \dots < i_k \leq n} \left( \prod_{j=1}^k \epsilon_{i_j} - \prod_{j=1}^k \epsilon'_{i_j} \right) \prod_{j=1}^k \lambda_{i_j} = 0. \end{aligned} \quad (3.54)$$

Remarquons alors que pour  $k$  pair, on peut choisir  $\epsilon_i = 1$  et  $\epsilon'_i = -1$  pour tout  $i \in \{1, \dots, n\}$  de telle sorte que :

$$\prod_{j=1}^k \epsilon_{i_j} - \prod_{j=1}^k \epsilon'_{i_j} = 0, \quad 1 \leq i_1 < \dots < i_k \leq n. \quad (3.55)$$

Ainsi, (3.53) n'est pas vérifiée pour  $k$  pair et  $b_{n-k}$  n'est pas séparant pour  $V(\langle \mathcal{B} \rangle)$ .

Supposons maintenant  $k$  impair. Tout d'abord, considérons l'égalité suivante :

$$\prod_{j=1}^k \epsilon_{i_j} - \prod_{j=1}^k \epsilon'_{i_j} = 0, \quad 1 \leq i_1 < \dots < i_k \leq n. \quad (3.56)$$

Posons :

$$\epsilon_i = (-1)^{p_i}, \quad p_i \in \{0, 1\}, \quad 1 \leq i \leq n.$$

L'égalité (3.56) se réécrit alors :

$$1 \leq i_1 < \dots < i_k \leq n, \quad \prod_{j=1}^k \epsilon_{i_j} - \prod_{j=1}^k \epsilon'_{i_j} = 0 \quad \Leftrightarrow \quad (-1)^{\sum_{j=1}^k p_{i_j}} = (-1)^{\sum_{j=1}^k p'_{i_j}}.$$

Ainsi, trouver deux couples  $\epsilon$  et  $\epsilon'$  tels que (3.56) soit vérifié peut s'écrire comme un problème d'injectivité de l'application linéaire  $\phi$  suivante :

$$\begin{array}{ccc} \mathbb{F}_2^n & \xrightarrow{\phi} & \mathbb{F}_2^{\binom{n}{k}} \\ (p_1, \dots, p_n) & \mapsto & \left( \sum_{j=1}^k p_{i_j}, \dots, \sum_{j=1}^k p_{i_j} \right). \end{array}$$

Nous notons  $A_{n,k}$  sa matrice associée dans les bases canoniques de  $\mathbb{F}_2^n$  et  $\mathbb{F}_2^{\binom{n}{k}}$ . Celle-ci est formée de vecteurs lignes de longueur  $n$  contenant  $k$  entrées à 1 et  $n-k$  entrées à zéro. Aussi, la variable  $b_{n-k}$  est séparante si et seulement si le rang de la matrice  $A_{n,k}$  est égal à  $n$ .

Dans un premier temps, nous pouvons remarquer que l'on peut extraire une sous-matrice circulante de la matrice  $A_{n,k}$ . Celle-ci a la forme suivante :

$$A_s := \begin{pmatrix} 1 & \dots & \dots & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 & \dots & \dots & 1 \\ 1 & \ddots & & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Notons par  $h$  le polynôme ayant pour coefficients les coefficients de la dernière ligne de  $A_s$ , c'est-à-dire :

$$h := x^{n-1} + \dots + x^{n-k+1} + 1.$$

D'après [Ing56], le rang de la matrice circulante  $A_s$  est égale au degré du plus grand diviseur commun de  $x^n - 1$  et  $h$ . Comme le nombre de monômes de  $h$  (c'est-à-dire  $k$ ) est impair, alors  $h$  est irréductible sur  $\mathbb{F}_2$ . Ainsi, le plus grand diviseur entre  $x^n - 1$  et  $h$  est trivial. Ainsi, le rang de la matrice circulante  $A_s$  est  $n$ . Il en est donc de même pour  $A_{n,k}$ . Ainsi, l'égalité (3.56) n'est vérifiée que pour  $\epsilon = \epsilon'$ . Nous dirons alors que  $b_{n-k}$  est séparante génériquement pour  $k$  impair.

Cherchons maintenant à caractériser la condition pour laquelle (3.56) est vérifiée lorsque  $\epsilon \neq \epsilon'$ , condition de *non-généricité* de la séparabilité de  $b_{n-k}$  pour  $k$  impair. Soit  $\beta_{(i_1, \dots, i_k)} \in \{-2, 0, 2\}$ ,  $1 \leq i_1 < \dots < i_k \leq n$ . La condition de généricité pour laquelle la variable  $b_{n-k}$  est séparante pour  $k$  impair est :

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \beta_{(i_1, \dots, i_k)} \prod_{j=1}^k \lambda_{i_j} = 0, \quad \Rightarrow \quad \forall i_j, \quad i_1 < \dots < i_k, \quad \beta_{(i_1, \dots, i_k)} = 0.$$

En d'autres termes,  $b_{n-k}$  n'est pas séparante pour  $k$  impair s'il existe  $\beta_{(i_1, \dots, i_k)} \in \{-2, 0, 2\}$ ,  $1 \leq i_1 < \dots < i_k \leq n$  non tous triviaux (c'est-à-dire non tous zéro) tels que :

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \beta_{(i_1, \dots, i_k)} \prod_{j=1}^k \lambda_{i_j} = 0. \quad (3.57)$$

Par ailleurs, les racines  $\epsilon_j \lambda_j$  de  $f$  sont en relation avec les coefficients de son facteur spectral  $g$  via les identités de Newton :

$$k = 1, \dots, n, \quad d_{2(n-k)} = \sum_{\substack{1 \leq i_1, \dots, i_k \leq n \\ i_1 < \dots < i_k}} \prod_{j=1}^k \epsilon_{i_j} \lambda_{i_j}. \quad (3.58)$$

Considérons le système algébrique formé par les  $n$  équations (3.58) et la condition de non-généricité (3.57). Ce système comprend  $n+1$  équations dans les  $2n$  variables  $\{\lambda_j, d_{2(n-j)}\}_{j=1, \dots, n}$ . On peut alors éliminer les variables  $\lambda_j$  de manière à obtenir un polynôme  $D_k \in \mathbb{Q}[d_0, \dots, d_{2(n-1)}]$  s'annulant si et seulement si le polynôme  $f$  n'est pas générique par rapport à la variable  $b_{n-k}$ , ce qui conclut la preuve.  $\square$

Ainsi, *génériquement*,  $\langle \mathcal{B} \rangle$  est radical, et, par exemple,  $b_{n-1}$  est un élément séparant  $\mathcal{B}$  (le choix de  $b_{n-1}$ , plutôt qu'un autre  $b_{n-k}$ , comme élément séparant dans le cas générique sera motivé dans la Section 3.5.4). Cherchons maintenant à déterminer une paramétrisation de  $V(\langle \mathcal{B} \rangle)$ .

### 3.5.3 Paramétrisation de $V(\langle \mathcal{B} \rangle)$

$b_{n-1}$  séparant  $V(\langle \mathcal{B} \rangle)$ , nous pouvons obtenir la *base de Gröbner sous forme shape* suivante.

**Théorème 3.16.** Dans le cas générique,  $\mathcal{B}$  possède une *base de Gröbner sous forme shape* par rapport à un ordre d'élimination  $\{b_0, \dots, b_{n-2}\} \succ \succ b_{n-1}$ . Cette base a la forme suivante :

$$\mathcal{S} := \begin{cases} p(b_{n-1}) = 0, \\ b_{n-2} - p_{n-2}(b_{n-1}) = 0, \\ \vdots \\ b_0 - p_0(b_{n-1}) = 0. \end{cases}$$

$p$  est un polynôme univarié de degré  $2^n$  et les  $p_i$  sont des polynômes univariés de degré strictement inférieur à  $2^n$ .

*Démonstration.* Ce théorème découle immédiatement de la propriété de séparabilité de  $b_{n-1}$  et de la dimension de  $\mathcal{A}$ .  $\square$

Pour obtenir une telle base de Gröbner, il existe un grand nombre d'algorithmes. En début de chapitre, nous présentons le plus basique d'entre-eux : l'*algorithme de Buchberger*. Cet algorithme permet d'obtenir une base de Gröbner d'un idéal polynomial en un nombre fini d'étapes, en calculant des réductions de  $S$ -polynômes. Une version raffinée de l'algorithme de Buchberger est donnée dans [CLO15, Chap. 2, §10]. Durant le déroulement de l'algorithme de Buchberger, on se retrouve souvent à calculer des réductions de  $S$ -polynômes triviales. Celles-ci peuvent être évitées afin de gagner en efficacité. C'est ce que cherche, par exemple, à faire l'algorithme  $F_5$  [Fau02], expliqué brièvement dans [CLO15, Chap. 10, §4]. Le lecteur intéressé est invité à lire [CLO15, Chap. 10] expliquant plusieurs algorithmes de calcul de

bases de Gröbner.

Tous ces algorithmes n'utilisent pas le fait que l'on part d'une base de Gröbner pour un certain ordre monomial. En effet, le système  $\mathcal{B}$  est déjà une base de Gröbner par rapport à l'ordre DRL  $b_{n-1} \succ \dots \succ b_0$ . Il existe d'autres algorithmes que ceux cités ci-dessus permettant de calculer une base de Gröbner pour un certain ordre monomial d'un idéal à partir d'une base de Gröbner pour un autre ordre monomial. C'est ce qu'on appelle les algorithmes de changement d'ordre. Le plus connu est l'*algorithme FGLM* [FGLM93]. Dans le cas du système  $\mathcal{B}$ , ces algorithmes de changement d'ordre sont bien plus efficaces que des algorithmes "classiques" pour calculer une base de Gröbner sous forme shape.

L'inconvénient majeur du calcul d'une base de Gröbner sous forme shape est la taille et le degré des polynômes qui la définissent. Reprenons les notations du Théorème 3.16. Dans le cas générique, le polynôme univarié  $p$  sera toujours de degré  $2^n$ . Il est donc impossible d'en trouver un de degré plus faible encodant l'ensemble des solutions. En revanche, les polynômes  $p_i$  sont en général de degré élevé (au plus  $2^n - 1$ ) et la taille de leurs coefficients grossissent en fonction du nombre de paramètres.

Une manière plus simple de paramétrer les zéros de  $\mathcal{B}$  est de considérer une *représentation rationnelle*. Plus particulièrement, nous proposons de calculer une *Représentation Univariée Rationnelle* (RUR) des zéros de  $\mathcal{B}$  [Rou99, Rou07]. Il s'agit d'une forme bijective entre  $V(\langle \mathcal{B} \rangle)$  et les racines d'un polynôme univarié  $p$  :

$$\begin{array}{ccc} V(\langle \mathcal{B} \rangle) & \longrightarrow & V(\langle p \rangle) \\ b = (b_0, \dots, b_{n-1}) & \longmapsto & \xi, \\ \left( \frac{q_0(\xi)}{q(\xi)}, \dots, \frac{q_{n-1}(\xi)}{q(\xi)} \right) & \longleftarrow & \xi. \end{array} \quad (3.59)$$

Une telle représentation peut-être calculée en utilisant l'algorithme donné dans [Rou07]. Cet algorithme ne requiert que la connaissance d'une base de Gröbner de  $\mathcal{A} = R/\langle \mathcal{B} \rangle$  et d'un algorithme de réduction calculant une forme normale modulo la base de Gröbner de l'idéal  $\langle \mathcal{B} \rangle$ .

**Théorème 3.17.** [Rou99] Le polynôme univarié  $p$  de (3.59) est le polynôme caractéristique de la matrice de multiplication  $M_{b_{n-1}}$  par  $b_{n-1}$  dans  $\mathcal{A}$  (voir Section 3.1.8, Définition 3.17), c'est-à-dire :

$$p := \det (M_{b_{n-1}} - \lambda I_{2^n}).$$

De plus, si  $\gcd(f, g)$  désigne le PGCD de deux polynômes  $f$  et  $g$ , alors notons

$$\bar{p}(b_{n-1}) := \frac{p}{\gcd\left(p, \frac{dp}{db_{n-1}}\right)} = \sum_{i=0}^d v_i b_{n-1}^{d-i} \in \mathbb{K}[b_{n-1}]$$

la *partie sans facteur carré* de  $p$  et définissons les polynômes suivants :

$$H_j(b_{n-1}) := \sum_{i=0}^j v_i b_{n-1}^{j-i} \in \mathbb{K}[b_{n-1}], \quad j = 0, \dots, d-1.$$

Alors, les polynômes  $p_i$  de (3.59) peuvent être obtenus de la manière suivante

$$\begin{cases} q(b_{n-1}) := \sum_{i=0}^{d-1} \text{Trace}(M_{b_{n-1}}^i) H_{d-i-1}(b_{n-1}), \\ q_k(b_{n-1}) := \sum_{i=0}^{d-1} \text{Trace}(M_{b_k} M_{b_{n-1}}^i) H_{d-i-1}(b_{n-1}), \quad k = 0, \dots, n-2, \end{cases}$$

où  $M_{b_i}$  désigne la matrice de multiplication par  $b_i$ .

**Exemple 3.14.** Soit  $n = 3$ . Supposons  $a \in \mathbb{Q}^n$  et  $c \in \mathbb{Q}^n$  tels que  $4d_2 - d_4^2 \neq 0$ . Une base de Gröbner de  $\mathcal{B}$  obtenue pour l'ordre lexicographique  $b_0 \succ b_1 \succ b_2$  donne la paramétrisation suivante :

$$\begin{cases} p = b_2^8 - 4d_4 b_2^6 + (6d_4^2 - 8d_2) b_2^4 + (-4d_4^3 + 16d_2 d_4 - 64d_0) b_2^2 + d_4^4 - 8d_2 d_4^2 + 16d_2^2, \\ b_1 = \frac{1}{2} b_2^2 - \frac{1}{2} d_4, \\ b_0 = \frac{1}{8} \frac{b_2(b_2^6 - 4d_4 b_2^4 + 5d_4^2 b_2^2 - 4d_2 b_2^2 - 2d_4^3 + 8d_2 d_4 - 64d_0)}{4d_2 - d_4^2}. \end{cases}$$

Quant à lui, le calcul d'une RUR donne le résultat suivant

$$\begin{cases} p = b_2^8 - 4d_4 b_2^6 + (6d_4^2 - 8d_2) b_2^4 + (-4d_4^3 + 16d_2 d_4 - 64d_0) b_2^2 + d_4^4 - 8d_2 d_4^2 + 16d_2^2, \\ q = 8b_2(b_2^6 - 3d_4 b_2^4 + 3d_4^2 b_2^2 - 4d_2 b_2^2 - d_4^3 + 4d_2 d_4 - 16d_0), \\ q_1 = 16b_2(d_2 b_2^4 - 2d_2 d_4 b_2^2 + 12d_0 b_2^2 + d_2 d_4^2 + 4d_0 d_4 - 4d_2^2), \\ q_0 = 16d_0(3b_2^4 - 2d_4 b_2^2 - d_4^2 + 4d_2), \end{cases}$$

c'est-à-dire :

$$\begin{cases} p = b_2^8 - 4d_4 b_2^6 + (6d_4^2 - 8d_2) b_2^4 + (-4d_4^3 + 16d_2 d_4 - 64d_0) b_2^2 + d_4^4 - 8d_2 d_4^2 + 16d_2^2, \\ b_1 = 2 \frac{d_2 b_2^4 - 2d_2 d_4 b_2^2 + 12d_0 b_2^2 + d_2 d_4^2 + 4d_0 d_4 - 4d_2^2}{b_2^6 - 3d_4 b_2^4 + 3d_4^2 b_2^2 - 4d_2 b_2^2 - d_4^3 + 4d_2 d_4 - 16d_0}, \\ b_0 = 2 \frac{d_0(3b_2^4 - 2d_4 b_2^2 - d_4^2 + 4d_2)}{b_2(b_2^6 - 3d_4 b_2^4 + 3d_4^2 b_2^2 - 4d_2 b_2^2 - d_4^3 + 4d_2 d_4 - 16d_0)}. \end{cases}$$

Cet exemple illustre que le calcul d'une RUR et celui d'une base de Gröbner pour l'ordre lexicographique donnent tous les deux une paramétrisation des solutions de  $\langle \mathcal{B} \rangle$ . En revanche, il contredit ce que nous disions plus haut sur la taille des objets, la RUR obtenue est visiblement moins compacte que la base de Gröbner. Ce résultat ne se généralise pas aux ordres plus élevés. Nous invitons le lecteur à effectuer quelques tests pour  $n = 4$  ou  $n = 5$  pour s'en rendre compte. Néanmoins, les paramétrisations obtenues sont très conséquentes dans le cas d'une RUR, et nous choisissons de ne pas les écrire ici.

Remarquons aussi que la condition  $4d_2 - d_4^2 = 0$  implique que  $b_2$  n'est plus séparante. Pour expliquer ceci, calculons le discriminant de  $p$  en  $b_2$  :

$$\text{disc}(p, b_2) = 128^8 d_0^4 (4d_2 - d_4^2)^2 (4d_0 d_4^3 - d_2^2 d_4^2 - 18d_0 d_2 d_4 + 4d_2^3 + 27d_0^2)^2.$$

Ainsi, si la condition  $4d_2 - d_4^2 = 0$  est vérifiée, alors le discriminant de  $p$  en  $b_2$  s'annule. Celui-ci possède alors au moins une racine de multiplicité 2 ou plus. Autrement dit, sous cette même condition, au moins deux points de  $V(\langle \mathcal{B} \rangle)$  ont la même coordonnée en  $b_2$ , ce qui veut dire que  $b_2$  n'est pas séparante. Remarquons aussi que cette condition intervient au dénominateur de la paramétrisation de  $b_0$  obtenue par un calcul de base de Gröbner pour l'ordre lexicographique  $b_0 \succ b_1 \succ b_2$ .

Calculer une RUR est intéressant en pratique puisque nous obtenons une paramétrisation des solutions  $X$  de  $\mathcal{R} = 0$  suivant la variable  $b_{n-1}$ , laquelle satisfait une “simple” équation polynomiale  $p(b_{n-1}) = 0$ . Il ne reste alors plus qu’à substituer les racines de  $p$  dans la matrice  $X(b_{n-1})$  afin d’obtenir les solutions de l’ERA  $\mathcal{R} = 0$ .

### 3.5.4 Caractérisation de $X > 0$ solution de $\mathcal{R} = 0$

Nous avons montré qu’il est intéressant de calculer une RUR de  $V(\langle \mathcal{B} \rangle)$  afin de paramétrer  $V(\langle \mathcal{B} \rangle)$  en fonction de la variable  $b_{n-1}$  :

$$\mathcal{R} := \begin{cases} p(b_{n-1}) = 0, \\ b_{n-2} = \frac{q_{n-2}(b_{n-1})}{q(b_{n-1})}, \\ \vdots \\ b_0 = \frac{q_0(b_{n-1})}{q(b_{n-1})}. \end{cases}$$

Cette paramétrisation, combinée avec le Théorème 3.11, permet de paramétrer les solutions  $X$  de  $\mathcal{R} = 0$  en fonction de la variable  $b_{n-1}$ . Génériquement, parmi ces  $2^n$  solutions possibles, il reste à caractériser celle définie positive.

Pour vérifier si une matrice est définie positive, on peut utiliser le *critère de Sylvester* [Gil91]. Rappelons ce critère.

**Définition 3.23.** Soit  $M \in \mathbb{K}^{n \times n}$  une matrice symétrique réelle dont les coefficients sont notées par  $m_{i,j}$ , c’est-à-dire  $M = (m_{i,j})_{1 \leq i,j \leq n}$ . Les *mineurs principaux dominants* de  $M$  sont les déterminants des sous matrices  $M_k := (m_{i,j})_{1 \leq i,j \leq k}$  pour  $k = 1, \dots, n$ .

**Théorème 3.18.** (*Critère de Sylvester*) Une matrice symétrique réelle est définie positive si et seulement si ses mineurs principaux dominants sont positifs.

Disposant de la paramétrisation rationnelle  $\mathcal{R}$ , tous les coefficients de  $X$  se réécrivent alors rationnellement en fonction de  $b_{n-1}$ , où  $b_{n-1}$  est racine de  $p$ . Ainsi, tester la positivité des mineurs principaux dominants de  $X(b_{n-1})$ , notés par  $m_i$  pour  $i = 1, \dots, n$ , revient à tester la positivité de  $n$  fractions rationnelles sous la contrainte que  $b_{n-1}$  est racine de  $p$  :

$$\begin{cases} X(b_{n-1}) > 0, \\ p(b_{n-1}) = 0, \end{cases} \Leftrightarrow \begin{cases} m_i(b_{n-1}) > 0, & i = 1, \dots, n \\ p(b_{n-1}) = 0. \end{cases} \quad (3.60)$$

Ce critère, qui se ramène à un *semi-algébrique* (c’est-à-dire un systèmes d’équations et d’inéquations), est difficile à tester en pratique. [AHKY09] propose une alternative basée sur la factorisation spectrale permettant de transformer ce critère semi-algébrique en critère purement algébrique.

**Théorème 3.19.** [AHKY09] Soient  $a \in \mathbb{R}^n$  et  $c \in \mathbb{R}^n$  tels que  $G$  soit commandable et observable. Soit  $\sigma := \max\{b_{n-1} \in \mathbb{R} \mid p(b_{n-1}) = 0\}$ . La solution définie positive de  $\mathcal{R} = 0$  est alors donnée par  $X(\sigma)$ , c’est-à-dire  $X(\sigma) > 0$ .

*Démonstration.* La Proposition 3.5, vue au début de ce chapitre, montre comment construire une solution de  $\mathcal{R} = 0$  à partir de vecteurs propres de  $\mathcal{H}$ . En considérant  $n$  vecteurs propres associés à  $n$  valeurs propres à partie réelle négative, [ACA<sup>+</sup>99] montre que l’on peut obtenir

l'unique solution définie positive de  $\mathcal{R} = 0$ . Calculer ces valeurs propres à partie réelle négative revient à trouver un polynôme  $g(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i) = \sum_{i=1}^n b_i \lambda^i$ , où  $\lambda_i \in \mathbb{C}_-$ , tel que :

$$f(\lambda) = (-1)^n g(\lambda) g(-\lambda).$$

Un tel polynôme  $g$  est donc défini par une solution particulière du système polynomial  $\mathcal{E}$ . En particulier, nous avons :

$$b_{n-1} = - \sum_{i=1}^n \lambda_i, \quad \lambda_i \in \mathbb{C}_-.$$

Notons que dans le cas où  $g$  ne comporte pas seulement des valeurs propres à partie réelle négative, nous avons :

$$b_{n-1} = - \sum_{i=1}^n \epsilon_i \lambda_i, \quad \epsilon_i \in \{-1, 1\}, \quad \lambda_i \in \mathbb{C}_-.$$

Comme tous les  $\lambda_i$  sont à partie réelle négative,  $b_{n-1} = - \sum_{i=1}^n \lambda_i \in \mathbb{R}$  est maximal dans le sens où aucune autre combinaison  $-\sum_{i=1}^n \epsilon_i \lambda_i$  avec  $\epsilon_i \in \{-1, 1\}$  n'est supérieure à celui-ci.

De plus, génériquement, la variable  $b_{n-1}$  est séparante. Nous disposons alors d'un polynôme univarié  $p$  de degré  $2^n$  (par calcul du polynôme caractéristique de la matrice de multiplication par  $b_{n-1}$  dans  $\mathcal{A} = R/\langle \mathcal{B} \rangle$  par exemple) dont les racines sont les coordonnées  $b_{n-1}$  de chaque zéro de  $V(\langle \mathcal{B} \rangle)$ . Ainsi, la racine maximale de  $p$  définit la solution définie positive de l'ERA  $\mathcal{R} = 0$ .  $\square$

Enfin, la solution définie positive de  $\mathcal{R} = 0$  vérifie la propriété suivante.

**Proposition 3.11.** Soient  $a \in \mathbb{R}^n$  et  $c \in \mathbb{R}^n$  tels que  $G$  soit commandable et observable. Alors,  $X > 0$  vérifie  $b_0 = \sqrt{a_0^2 + c_0^2}$ .

*Démonstration.* De part le lien entre les  $b_i$  et le facteur spectral  $g$  de  $f$ , nous avons :

$$b_0 = (-1)^n \prod_{i=1}^n \lambda_i, \quad \lambda_i \in \mathbb{C}_-.$$

Rappelons que si  $\lambda_i$  est racine de  $g$ , alors  $\bar{\lambda}_i$  l'est aussi. De plus, nous avons  $\lambda_i \bar{\lambda}_i = |\lambda_i|^2 > 0$ . Notons alors que nous avons deux cas à étudier.

- Soit  $n$  pair. Supposons que  $g$  possède  $n - N$  racines réelles (nombre pair), notées  $\lambda_{N+1}, \dots, \lambda_n$ . Ses  $N$  racines complexes sont alors notées  $\lambda_1, \dots, \lambda_{N/2}$  et  $\bar{\lambda}_1, \dots, \bar{\lambda}_{N/2}$ . Alors, nous avons :

$$b_0 = (-1)^n \prod_{i=N+1}^n \lambda_i \prod_{i=1}^{N/2} \lambda_i \bar{\lambda}_i = \prod_{i=N+1}^n \lambda_i \prod_{i=1}^{N/2} |\lambda_i|^2.$$

Or  $\prod_{i=N+1}^n \lambda_i > 0$  car c'est un produit d'un nombre pair de  $\lambda_i < 0$ . De plus,  $\prod_{i=1}^{N/2} |\lambda_i|^2 > 0$ .  
D'où  $b_0 > 0$ .

- Soit  $n$  impair. Supposons que  $g$  possède  $n - N$  racines réelles (nombre impair), notées  $\lambda_{N+1}, \dots, \lambda_n$ . Ses  $N$  racines complexes sont alors notées  $\lambda_1, \dots, \lambda_{N/2}$  et  $\bar{\lambda}_1, \dots, \bar{\lambda}_{N/2}$ . Alors, nous avons :

$$b_0 = (-1)^n \prod_{i=N+1}^n \lambda_i \prod_{i=1}^{N/2} \lambda_i \bar{\lambda}_i = (-1)^n \prod_{i=N+1}^n \lambda_i \prod_{i=1}^{N/2} |\lambda_i|^2.$$

Or  $\prod_{i=N+1}^n \lambda_i < 0$  car c'est un produit d'un nombre impair de  $\lambda_i < 0$ . De plus, nous avons  $\prod_{i=1}^{N/2} |\lambda_i|^2 > 0$ . D'où, nous obtenons  $b_0 > 0$ .

Ainsi, nous obtenons  $b_0 > 0$  quel que soit l'ordre  $n$  du système. De plus, de  $\mathcal{B}_0 = 0$  (voir (3.43) et (3.42) pour  $k = 0$ ), nous déduisons  $b_0 = \pm \sqrt{a_0^2 + c_0^2}$ . Ainsi, nous obtenons  $b_0 = \sqrt{a_0^2 + c_0^2}$ , ce qui conclut la preuve.  $\square$

Cette propriété permet de réduire la taille des objets que l'on étudie d'un facteur 2 lorsque l'on cherche uniquement à trouver la solution définie positive de  $\mathcal{R} = 0$ . En effet, en considérant  $b_0$  fixé par les paramètres du système  $G$ , c'est-à-dire

$$b_0 = \sqrt{a_0^2 + c_0^2},$$

nous pouvons alors réduire le système polynomial  $\mathcal{E}$  en retirant l'équation polynôme  $\mathcal{B}_0 = 0$  :

$$\begin{aligned} \tilde{\mathcal{B}} &:= \{\mathcal{B}_k\}_{k=1..n-1}, \\ \tilde{\mathcal{E}} &:= \{\mathcal{B}_k = 0\}_{k=1..n-1}. \end{aligned} \quad (3.61)$$

$b_0$  devient alors une valeur fixe du système polynomial  $\tilde{\mathcal{B}}$ , qui ne possède alors plus que  $n - 1$  inconnues  $b_k$ . Nous pouvons alors déduire une paramétrisation de ce système en fonction de  $b_{n-1}$  qui sera de taille plus faible que la précédente (facteur 2). Par exemple, le calcul d'une RUR de  $V(\langle \tilde{\mathcal{B}} \rangle)$  donne la paramétrisation suivante :

$$\tilde{\mathcal{R}} := \begin{cases} \tilde{p}(b_{n-1}, b_0) = 0, \\ b_{n-2} = \frac{\tilde{q}_{n-2}(b_{n-1}, b_0)}{\tilde{q}(b_{n-1}, b_0)}, \\ \vdots \\ b_1 = \frac{\tilde{q}_1(b_{n-1}, b_0)}{\tilde{q}(b_{n-1}, b_0)}, \\ b_0 = \sqrt{a_0^2 + c_0^2}. \end{cases} \quad (3.62)$$

Le degré de  $\tilde{p}$  est alors  $2^{n-1}$ , et les degrés des autres polynômes de la RUR  $\tilde{\mathcal{B}}_r$  sont inférieurs à  $2^{n-1}$ . La taille des objets étudiés est donc divisée par 2.

En pratique, ceci est très intéressant car les calculs sont fortement limités par l'ordre du système. Cette propriété nous permet de travailler sur des système d'un ordre  $n + 1$ , ce qui est non négligeable. Par ailleurs, ceci est très avantageux dans le cas où l'on cherche à trouver des expressions sous forme close des solutions de  $\mathcal{R} = 0$  comme nous allons le voir dans la Section 3.6.2.

### 3.5.5 Isolation certifiée des zéros de $\mathcal{B}$

Un des avantages de l'approche symbolique par rapport à des approches numériques classiques pour résoudre une équation de Riccati algébrique est de pouvoir *certifier numériquement* tout résultat. Une approximation certifiée des solutions d'un système d'équations est une représentation des solutions de ce système sous forme d'intervalles contenant les solutions réelles du système. Considérons, par exemple, l'équation  $p(b_{n-1}) = 0$ . Une approximation certifiée des racines de  $p$  donnera des intervalles pour les valeurs de  $b_{n-1}$  contenant à *coup sûr* les racines de  $p$ . Pour plus de détails sur l'analyse par intervalles, nous nous référons, par exemple, à [Jau01]. Pour des applications de l'analyse par intervalles, citons, par exemple, [DLDJ<sup>+</sup>07].

Il existe de multiples algorithmes réalisant une isolation certifiée des racines de  $p$ . Citons, par exemple, [BF00] détaillant un algorithme se basant sur la *méthode d'Aberth*, ou encore [RZ04] dont l'algorithme se base sur la *règle des signes de Descartes* [CA76]. La commande *Maple RootFinding/Isolate* implémente cette dernière méthode.

Les méthodes numériques classiques de calcul de solutions d'équations de Riccati algébriques (voir, par exemple, [BG96] pour une liste non-exhaustive de telles méthodes) ne sont pas certifiées numériquement. L'avantage de disposer d'une RUR de  $V(\mathcal{B})$  paramétrant les solutions de l'équation de Riccati algébrique est de pouvoir utiliser ces algorithmes de calcul certifié. En effet, la substitution des racines de  $p$  sous forme d'intervalles dans la RUR  $\mathcal{R}$  permet d'obtenir des intervalles certifiés pour les  $b_k$  solutions de  $\mathcal{E}$ .

**Exemple 3.15.** Soit  $n = 3$ . Supposons que les valeurs des paramètres sont fixés aux valeurs suivantes :

$$a_0 := 0, \quad c_0 := 10, \quad a_1 := 1, \quad c_1 := 2, \quad c_2 := 0, \quad a_2 := 5.$$

Le calcul d'une RUR de  $V(\mathcal{B})$  pour ces valeurs de paramètres donne le résultat suivant :

$$\begin{cases} p = b_2^8 - 92 b_2^6 + 3134 b_2^4 - 53228 b_2^2 + 259081, \\ b_1 = \frac{10 (b_2^4 + 194 b_2^2 + 2349)}{b_2^6 - 69 b_2^4 + 1567 b_2^2 - 13307}, \\ b_0 = \frac{200 (3 b_2^4 - 46 b_2^2 - 509)}{b_2 (b_2^6 - 69 b_2^4 + 1567 b_2^2 - 13307)}. \end{cases}$$

L'isolation des racines réelles de  $p$  via la commande *RootFinding/Isolate* de *Maple* avec 10 digits donne les intervalles exacts suivants, et leur évaluation avec 20 chiffres significatifs donne les intervalles approximés suivants :

$$\begin{aligned} \left[ -\frac{940430083379}{137438953472}, -\frac{470215041689}{68719476736} \right] &\approx [-6.8425294257685891353, -6.8425294257613131776], \\ \left[ -\frac{377030844431}{137438953472}, -\frac{188515422215}{68719476736} \right] &\approx [-2.7432604433197411709, -2.7432604433124652132], \\ \left[ \frac{188515422215}{68719476736}, \frac{377030844431}{137438953472} \right] &\approx [2.7432604433124652132, 2.7432604433197411709], \\ \left[ \frac{470215041689}{68719476736}, \frac{940430083379}{137438953472} \right] &\approx [6.8425294257613131776, 6.8425294257685891353]. \end{aligned}$$

La substitution respective de ces intervalles dans la RUR calculée précédemment donne les

intervalles respectifs suivants pour  $b_0$

$$\begin{aligned} \left[ -\frac{351843720888325}{35184372088832}, -\frac{351843720888309}{35184372088832} \right] &\approx [-10.000000000000142109, -9.9999999999996873612], \\ \left[ -\frac{351843720888321}{35184372088832}, -\frac{351843720888319}{35184372088832} \right] &\approx [-10.000000000000028422, -9.999999999999715783], \\ \left[ \frac{351843720888319}{35184372088832}, \frac{351843720888321}{35184372088832} \right] &\approx [9.999999999999715783, 10.000000000000028422], \\ \left[ \frac{351843720888309}{35184372088832}, \frac{351843720888325}{35184372088832} \right] &\approx [9.9999999999996873612, 10.000000000000142109], \end{aligned}$$

et les intervalles respectifs suivants pour  $b_1$  :

$$\begin{aligned} \left[ \frac{419049547332085}{35184372088832}, \frac{209524773666047}{17592186044416} \right] &\approx [11.910104471214850719, 11.910104471215106514], \\ \left[ -\frac{136115336218805}{17592186044416}, -\frac{544461344875217}{70368744177664} \right] &\approx [-7.7372610700652444393, -7.7372610700652018068], \\ \left[ -\frac{136115336218805}{17592186044416}, -\frac{544461344875217}{70368744177664} \right] &\approx [-7.7372610700652444393, -7.7372610700652018068], \\ \left[ \frac{419049547332085}{35184372088832}, \frac{209524773666047}{17592186044416} \right] &\approx [11.910104471214850719, 11.910104471215106514]. \end{aligned}$$

**Remarque 3.20.** En pratique, l'isolation certifiée des racines de  $V(\langle \mathcal{B} \rangle)$  par un calcul de RUR est limité par l'ordre du système  $n$ . Par exemple, pour  $n = 8$ ,  $p$  est de degré 256. Il faut alors isoler les racines de  $p$  par intervalles avec une précision de 250 digits pour assurer une stabilité suffisante de la RUR. Pour plus de détails sur le degré des polynômes de la RUR et la taille de leurs coefficients, voir l'Annexe E. En pratique, à paramètres fixés, nous pouvons calculer une RUR de  $V(\mathcal{B})$  jusqu'à  $n = 10$  sur une machine personnelle en quelques minutes.

**Remarque 3.21.** Disposant d'une isolation certifiée des coefficients  $b_i$  de la factorisation spectrale  $g$  de  $f$ , alors en utilisant la Proposition 3.8 nous pouvons déduire une "factorisation copremière de  $G$  par intervalles".

## 3.6 Étude de $\mathcal{E}$ à paramètres non fixés

### 3.6.1 Remarques préliminaires

Supposons maintenant que les paramètres  $a$  et  $c$  ne sont pas fixés à des valeurs numériques. Nous aimerions étendre les résultats présentés aux sections précédentes à ce cas paramétrique.

Nous pouvons alors réutiliser les résultats du Théorème 3.11 donnant les coefficients de  $X$  solution de  $\mathcal{R} = 0$  en fonction des variables  $b_k$ . Disposant du système  $\mathcal{E}$ , nous aimerions trouver une paramétrisation de ses solutions. Cependant, celles-ci sont dans la clôture algébrique  $\overline{\mathbb{Q}(a, c)}$ , et non plus dans  $\mathbb{C}$ . Ainsi, la plupart des démonstrations vues précédemment ne sont plus valides car nous ne pouvons plus parler de  $\lambda_i \in \mathbb{C}_-$ .

Nous proposons alors une alternative, que nous développons ici, pour des systèmes d'ordre 4 au plus. Pour de tels systèmes, nous exhibons une autre représentation de  $\mathcal{E}$  par des opérations simples, et nous utilisons le concept de variété discriminante afin de décomposer l'espace des paramètres en cellules régulières dans lesquelles nous pouvons réutiliser les résultats de la section précédente.

Rappelons que nous notons  $\mathbb{K} := \mathbb{Q}(d_0, \dots, d_{2n})$ ,  $R := \mathbb{K}[b_0, \dots, b_{n-1}]$  et la  $\mathbb{K}$ -algèbre  $\mathcal{A} := R/\langle \mathcal{B} \rangle$ .

### 3.6.2 Paramétrisation des solutions de $\mathcal{E}$ pour $1 \leq n \leq 4$

Soit  $n \leq 4$ . Le Théorème 3.11 (voir page 91) montre que les coefficients de  $X$  solution de  $\mathcal{R} = 0$  sont fonction des variables  $b_k$ . La Table 3.1 explicite ces matrices  $X$  pour les ordres faibles considérés. Ces mêmes variables  $b_k$  satisfont le système d'équations polynomiales  $\mathcal{E}$  défini par  $\mathcal{B}$ , rappelé en Table 3.2 pour  $n \leq 4$ , et où les  $d_{2k}$  sont définis en Table 3.3.

**Remarque 3.22.** A la vue de ces Tables, nous remarquons que la complexité de la représentation des solutions grandit en fonction de  $n$  et du nombre de paramètres. Cet aspect est traité en Annexe E pour  $n$  quelconque.

D'après l'équation  $\mathcal{B}_0 = 0$ , la variable  $b_0$  est déterminée en fonction des paramètres  $a_0$  et  $c_0$ . Le cas  $n = 1$  est alors un cas particulier. Pour  $n \geq 2$ ,  $b_0$  peut être vu comme un nouveau paramètre comme nous l'avons vu à la section précédente. Nous étudions alors l'ensemble  $\mathcal{B}$  privé du premier polynôme  $\mathcal{B}_0$ , que l'on note par  $\tilde{\mathcal{B}}$ , de variables  $(b_1, \dots, b_{n-1})$  et de paramètres  $(b_0, d_2, \dots, d_{2n})$ . A cet ensemble, nous associons le système d'équations polynomiales défini par  $\tilde{\mathcal{E}} := \{\mathcal{B}_k = 0\}_{k=1, \dots, n-1}$ .

Ensuite, nous proposons de calculer une paramétrisation des solutions de  $\tilde{\mathcal{E}}$ . Par des opérations simples, nous montrons que ce système peut se réécrire sous une forme plus exploitable.

**Proposition 3.12.** Soit  $2 \leq n \leq 4$ . Soit  $\mathcal{B}$  un système polynomial défini dans la Table 3.2. Une paramétrisation rationnelle des solutions de  $\tilde{\mathcal{E}}$  est donnée par  $\tilde{\mathcal{R}}$  défini dans la Table 3.4.

*Démonstration.* • Soit  $n = 2$ . Le polynôme  $\mathcal{B}_1$  peut être considéré comme un polynôme univarié en la variable  $b_1$ . La paramétrisation obtenue est alors triviale.

- Soit  $n = 3$ . En utilisant  $\mathcal{B}_2 = 0$ , nous pouvons exprimer  $b_1$  en fonction de  $b_2$  :

$$b_1 = \frac{1}{2} (b_2^2 - d_4).$$

En substituant cette paramétrisation de  $b_1$  dans  $\mathcal{B}_1 = 0$ , nous obtenons le polynôme univarié  $\tilde{p}$  en  $b_2$  suivant :

$$\tilde{p} := b_2^4 - 2 d_4 b_2^2 - 8 b_0 b_2 + d_4^2 - 4 d_2 = 0.$$

Nous avons alors la paramétrisation rationnelle de  $\tilde{\mathcal{R}}$  suivante :

$$\tilde{\mathcal{R}} := \begin{cases} b_1 = \frac{1}{2} (b_2^2 - d_4), \\ b_2 = b_2, \\ \tilde{p} := b_2^4 - 2 d_4 b_2^2 - 8 b_0 b_2 + d_4^2 - 4 d_2 = 0. \end{cases} \quad (3.63)$$

- Soit  $n = 4$ . En utilisant  $\mathcal{B}_3 = 0$ , nous pouvons exprimer  $b_2$  en fonction de  $b_3$  :

$$b_2 = \frac{1}{2} (b_3^2 - d_6). \quad (3.64)$$

En substituant cette paramétrisation de  $b_2$  dans  $\mathcal{B}_2 = 0$ , nous obtenons :

$$8 b_3 b_1 = b_3^4 - 2 d_6 b_3^2 + 8 b_0 + d_6^2 - 4 d_4. \quad (3.65)$$

Nous voyons une singularité apparaître pour  $b_3 = 0$ . Supposons alors, dans un premier temps,  $b_3 = 0$ . Alors, (3.65) devient :

$$8b_0 + d_6^2 - 4d_4 = 0. \quad (3.66)$$

De plus, (3.64) se simplifie en :

$$b_2 = \frac{-d_6}{2}. \quad (3.67)$$

La substitution de (3.67) dans  $\mathcal{B}_1$  donne :

$$b_1^2 + b_0 d_6 - d_2 = 0. \quad (3.68)$$

Ainsi,  $b_1$  est complètement déterminé par les valeurs des paramètres en utilisant l'équation (3.68). Les solutions de  $\tilde{E}$  sont alors données par :

$$\begin{cases} b_1^2 + b_0 d_6 - d_2 = 0, \\ b_2 = \frac{-d_6}{2}, \\ b_3 = 0. \end{cases}$$

Supposons maintenant que  $b_3 \neq 0$ . De (3.65), nous déduisons :

$$b_1 = \frac{b_3^4 - 2d_6 b_3^2 + 8b_0 + d_6^2 - 4d_4}{8b_3}. \quad (3.69)$$

Notons que nous retrouvons la singularité  $8b_0 + d_6^2 - 4d_4$  obtenues en (3.66). En substituant (3.69) et (3.64) dans  $\mathcal{B}_1 = 0$ , nous obtenons un polynôme univarié  $p$  en la variable  $b_3$  :

$$\begin{aligned} \tilde{p} := & b_3^8 - 4d_6 b_3^6 + (6d_6^2 - 8d_4 - 48b_0) b_3^4 \\ & + 4(-d_6^3 + 4(d_4 + 2b_0)d_6 - 16d_2) b_3^2 + (d_6^2 - 4d_4 + 8b_0)^2. \end{aligned} \quad (3.70)$$

Par conséquent, nous obtenons la paramétrisation rationnelle de  $\mathcal{B}$  suivante :

$$\tilde{\mathcal{R}} := \begin{cases} b_1 = \begin{cases} \pm\sqrt{d_2 - b_0 d_6} & \text{si } b_3 = 0, \\ \frac{b_3^4 - 2d_6 b_3^2 + 8b_0 + d_6^2 - 4d_4}{8b_3} & \text{sinon,} \end{cases} \\ b_2 = \frac{1}{2}(b_3^2 - d_6), \\ b_3 = b_3, \\ \tilde{p} = 0. \end{cases} \quad (3.71)$$

□

**Remarque 3.23.** Notons aussi que, disposant d'une paramétrisation des zéros de  $\langle \mathcal{B} \rangle$  pour des paramètres indéterminés, alors l'évaluation de ces paramètres en des valeurs numériques donne le même résultat que si nous calculions une paramétrisation des zéros de  $\langle \mathcal{B} \rangle$  particulier en ces mêmes valeurs numériques. Nous parlons alors de la propriété de *commutativité des spécialisations* [BPR03, Prop. 8.74]. En d'autres termes, soient  $d^* := (d_0^*, \dots, d_{2n}^*)$  des valeurs particulières de paramètres, alors calculer une paramétrisation de  $V(\langle \mathcal{B} \rangle)$  puis la particulariser en  $d^*$  est équivalent à particulariser  $\langle \mathcal{B} \rangle$  en  $d^*$  puis en calculer une paramétrisation de ses zéros.

**Remarque 3.24.** Il est possible d'étendre le calcul de telles paramétrisations à des ordres plus élevés. Cependant, ces calculs deviennent très coûteux lorsque le nombre de paramètres non-fixés  $a_i$  et  $c_i$  augmente, ou lorsque l'ordre  $n$  augmente. Dans cette optique, l'Annexe E explicite des ordres de grandeur des polynômes d'une RUR de  $V(\langle \mathcal{B} \rangle)$  à paramètres fixés. En pratique, nous ne pouvons calculer une paramétrisation de  $V(\langle \mathcal{B} \rangle)$ , où tous les  $a$  et  $c$  sont non nuls et non fixés, pour  $n = 6$  ou plus sur nos machines disposant de 16 GB de RAM, cette dernière saturant.

### 3.6.3 Solutions de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$

Des paramétrisations  $\tilde{\mathcal{R}}$  obtenues, notons que :

- pour  $n = 2$ ,  $\tilde{p}$  est de degré 2 en  $b_1$ ,
- pour  $n = 3$ ,  $\tilde{p}$  est de degré 4 en  $b_2$ ,
- pour  $n = 4$ ,  $\tilde{p}$  est de degré 4 en  $b_3^2$ .

Nous pouvons alors appliquer les résultats de l'Annexe D donnant des formules explicites des racines de polynômes à coefficients quelconques et de degré inférieur ou égal à 4. Ces racines par radicaux font intervenir des puissances non entières de nombres complexes. Afin de garder une unicité des notations, nous adoptons les notations suivantes

$$x \in \mathbb{C}, n \in \mathbb{Q}, \quad x^{1/n} = \exp\left(\frac{1}{n} (\ln(|x|) + i \arg(x))\right)$$

où  $\arg(x)$  est l'argument de  $x$  tel que  $-\pi \leq \arg(x) \leq \pi$ . Dans la littérature,  $x^{1/n}$  définie ci-dessus est appelée *racine  $n^{\text{ème}}$  principale* de  $x$ . Voir [BC09, p.26] pour plus de détails. Notons par ailleurs que les termes de tête des polynômes univariés de  $\tilde{\mathcal{R}}$  sont tous égaux à 1, ce qui simplifie l'expression des racines par radicaux de  $\tilde{p}$ . Celles-ci sont données dans la Table 3.5.

**Remarque 3.25.** Pour  $n = 3$ , le coefficient de  $\tilde{p}$  en  $b_2^2$ , c'est-à-dire  $\mu_1 := -8b_0$ , peut-être considéré comme non nul. En effet, son annulation correspond à  $b_0 = 0$ , et donc à  $a_0 = c_0 = 0$ , ce qui correspond à une dégénérescence d'ordre de  $G$ . Cette remarque permet d'éviter une variété singulière sur lesquelles les racines par radicaux ne sont pas définies.

**Remarque 3.26.** Notons aussi que, disposant d'une paramétrisation de  $V(\langle \mathcal{B} \rangle)$  pour  $n \geq 5$ , le polynôme univarié  $\tilde{p}$  obtenu est généralement de degré trop élevé pour expliciter ses racines par radicaux. Ainsi, l'utilisation de paramètres non fixés s'avère limité. C'est pourquoi nous choisissons de nous limiter à l'ordre au plus 4 dans cette section.

De ces racines exprimées par radicaux, l'utilisation de la paramétrisation  $\tilde{\mathcal{R}}$  et de la Table 3.1 permet de construire les matrices  $X$  solutions de  $\mathcal{R} = 0$ . Parmi ces expressions sous forme close, il reste à déterminer l'unique solution définie positive.

### 3.6.4 Solution définie positive de $\mathcal{R} = 0$ pour $1 \leq n \leq 4$

Avec la paramétrisation  $\tilde{\mathcal{R}}$  et le Théorème 3.11, nous disposons de  $X(a, c, b_0, b_{n-1})$ , où :

$$\begin{cases} \tilde{p}(b_{n-1}, b_0) = 0, \\ b_0^2 - d_0 = 0. \end{cases}$$

$b_0$  étant complètement déterminé par  $d_0 = a_0^2 + c_0^2$ , nous rappelons que nous le considérons comme un paramètre fixé.

Par ailleurs, nous disposons de formules par radicaux des racines de  $\tilde{p}$  (voir la Table 3.5). En particulier, cette table montre que les racines de  $\tilde{p}$  sont exprimables par une unique expression par radicaux sauf pour certaines valeurs de paramètres. Nous adoptons alors la notation suivante.

**Définition 3.24.** Nous notons par  $\mathcal{V}_s$  les variétés singulières sur lesquelles les expressions par radicaux des racines de  $\tilde{p}$  changent de forme.

Pour  $1 \leq n \leq 4$ , ces variétés singulières sont données en Table 3.6.

**Remarque 3.27.** Suivant la représentation par radicaux choisie, nous pouvons introduire différentes variétés singulières  $\mathcal{V}_s$ . Nous nous référons à [Mut80, Tig15, Kin96, Neu14, Shm11, Nic09, Hol02] ou encore [Zwi02] pour différentes méthodes de recherche de racines d'un polynôme sous formes closes. Dans ce mémoire, nous considérons uniquement les méthodes popularisées par Lagrange et Descartes, qui sont rappelées en Annexe D.

Il reste alors à identifier l'unique matrice définie positive.

Pour cela, nous pouvons appliquer le critère de Sylvester (Théorème 3.18), énoncé à la section précédente. Notant  $m_i$  les mineurs principaux dominants de  $X(a, c, b_{n-1})$ , celui-ci consiste à résoudre le système :

$$\begin{cases} m_i(a, c, b_0, b_{n-1}) > 0, & 1 \leq i \leq n, \\ \tilde{p} = 0, \\ b_0^2 - d_0 = 0. \end{cases} \quad (3.72)$$

Notons que  $m_i(a, c, b_0, b_{n-1}) \in \mathbb{Q}(a, c, b_0, b_{n-1})$ . Ainsi,  $a$  et  $c$  étant indéterminés, résoudre (3.72) consiste à déterminer le signe d'expressions rationnelles sous contrainte. En particulier, ces signes peuvent varier suivant les valeurs des paramètres  $a$  et  $c$ . Pour connaître les variations de signe de chaque  $m_i$  suivant les valeurs des paramètres  $a$  et  $c$ , il est d'abord nécessaire de caractériser son lieu d'annulation. Considérons le système polynomial obtenu en considérant les produits des numérateurs et dénominateurs de (3.72). Afin de caractériser le lieu d'annulation de ce système, nous pouvons en calculer une *Décomposition Cylindrique Algébrique* (CAD) [ACM84]. Une CAD est une décomposition en semi-algébriques connectés de  $\mathbb{R}^{2n}$  nommés cellules, dans lesquelles chaque polynôme a un signe constant. Cependant, cet outil n'est généralement pas efficace pour un nombre de paramètres supérieur à 2, à cause de la complexité doublement exponentielle de l'algorithme sous-jacent. Dans le cas du système (3.72), le nombre de paramètres est (au plus)  $2n$ . Ainsi, sauf pour des systèmes très particuliers à faible nombre de paramètres (certains  $a_i$  et  $c_i$  étant nuls), il apparaît difficile d'utiliser un tel critère semi-algébrique. Nous proposons alors une alternative.

Pour cela, nous proposons de calculer la variété discriminante de  $\tilde{\mathcal{R}}$ , c'est-à-dire celle de  $\tilde{p}$ . Ce polynôme univarié  $\tilde{p}$  est de la forme suivante

$$\tilde{p} = b_{n-1}^N + \theta_{N-1} b_{n-1}^{N-1} + \cdots + \theta_1 b_{n-1} + \theta_0, \quad (3.73)$$

où  $N = 2^{n-1}$  (voir Table 3.4). Notons que son terme de tête est égal à 1. Ainsi, la variété discriminante de  $\tilde{p}$  correspond exactement au discriminant  $\Delta_{\tilde{p}}$  de  $\tilde{p}$ , c'est-à-dire au lieu des paramètres où l'une des racines de  $\tilde{p}$  devient multiple. Nous avons alors :

$$\mathcal{V}_{\tilde{p}} := \{\theta \mid \Delta_{\tilde{p}}(\theta) = 0\}. \quad (3.74)$$

Suivant l'ordre  $n$  du système  $G$ , nous notons alors :

$$\mathcal{V}_n := \mathcal{V}_{\tilde{p}} \cup \mathcal{V}_s.$$

**Remarque 3.28.**  $\mathcal{V}_n$  est un sous-espace de l'espace des paramètres défini par  $a$ ,  $c$  et  $b_0$ .

D'un point de vue géométrique,  $\mathcal{V}_n$  décompose l'espace des paramètres en cellules régulières (c'est-à-dire des ouverts de l'espace des paramètres ne contenant pas  $\mathcal{V}_n$ ) dans lesquelles :

- le nombre de racines réelles de  $\tilde{p}$  est constant,
- les représentations des racines de  $\tilde{p}$  par radicaux sont identiques suivant les valeurs des paramètres  $a$ ,  $c$  et  $b_0$ .

Ces cellules régulières peuvent alors se caractériser suivant la même méthode qu'expliquée dans l'Exemple 3.9.

En tout point d'une cellule régulière, notée  $C_i$ , les racines par radicaux de  $\tilde{p}$  sont définies par la même expression par radicaux et sont de multiplicité 1. Par évaluation numérique de ces expressions en un point témoin  $(a, c, b_0)$  de  $C_i$ , nous pouvons identifier quelle expression identifie la racine réelle maximale de  $\tilde{p}$  dans  $C_i$ .

Par ailleurs, rappelons que la Proposition 3.11 montre que pour  $a$  et  $c$  fixés, alors  $X > 0$  solution de  $\mathcal{R} = 0$  vérifie  $b_0 > 0$ . Nous pouvons alors ne considérer que les cellules  $C_i$  dans lesquelles  $b_0$  prend des valeurs positives, les cellules pour lesquelles  $b_0$  est négatif ne nous permettant pas de distinguer la solution définie positive de  $\mathcal{R} = 0$ .

Ainsi, en particulierisant les racines par radicaux de la Table 3.5 en un point de chaque cellule  $C_i$  ne rencontrant pas  $\mathcal{V}_n$  et restreinte à  $b_0 > 0$ , nous identifions l'expression par radicaux représentant la racine réelle maximale  $\tilde{\sigma}$  de  $\tilde{p}$  dans cette cellule  $C_i$ . En particulierisant cette expression dans  $X(a, c, b_{n-1}, b_0)$  et en choisissant  $b_0 = \sqrt{a_0^2 + c_0^2}$ , nous obtenons une expression sous forme close de la solution définie positive  $X$  de  $\mathcal{R} = 0$  valable dans toute la cellule  $C_i$ .

En pratique, nous utilisons la commande *CellDecomposition* de *Maple* pour calculer ces cellules  $C_i$  restreintes à  $b_0 > 0$  et calculer un point témoin dans chacune d'elle. Le calcul des cellules  $C_i$  et l'identification des racines réelles maximales de  $\tilde{p}$  est résumé dans la proposition suivante.

**Proposition 3.13.** Soit  $n = 2$ . Alors la racine réelle maximale de  $\tilde{p}$  a toujours la même expression par radicaux et est définie par :

$$\tilde{\sigma} := b_1(1) = \sqrt{2b_0 + a_1^2 + c_1^2 - 2a_0}.$$

Soit  $n = 3$ . Alors, suivant les notations de la Table 3.5, la racine réelle maximale de  $\tilde{p}$  est :

$$\tilde{\sigma} := b_2(1, 1).$$

Rappelons que cette expression change de forme pour  $p_4 = 0$  (voir la Table 3.5).

Soit  $n = 4$ . La racine réelle maximale de  $\tilde{p}$  change d'expression suivant les valeurs des paramètres. Notons :

$$\left\{ \begin{array}{l} \mu_0 := 16(2b_0 - d_4)^2 - 64d_2d_6, \quad \mu_1 := -64(d_2 + b_0d_6), \quad \mu_2 := -8(6b_0 + d_4), \\ \Delta_Q := 256(\mu_0^3 + \beta_2\mu_0^2 + \beta_1\mu_0 + \beta_0), \\ \beta_2 := -\frac{\mu_2^2}{2}, \quad \beta_1 := \frac{\mu_2}{16}(\mu_2^3 + 9\mu_1^2), \quad \beta_0 := -\frac{\mu_1^2}{64}\left(\mu_2^3 + \frac{27}{4}\mu_1^2\right), \\ p_3 := -\frac{1}{3}\beta_2^2 + \beta_1, \quad q_3 := \frac{2}{27}\beta_2^3 - \frac{1}{3}\beta_1\beta_2 + \beta_0, \\ \delta_3 = 27(4p_3^3 + 27q_3^2), \quad \alpha_3 := \left(\frac{-27q_3 + \sqrt{\delta_3}}{2}\right)^{1/3}, \\ \forall i \in \{1, 2, 3\}, \quad r_i = \frac{1}{3}\left(\alpha_3 j^{2(i-1)} - \frac{3p_3}{\alpha_3 j^{(i-1)}}\right) - \frac{1}{3}\beta_2, \\ \mu_1^* := \sqrt{\frac{-8\mu_2^3}{27}}, \quad \mu_0^* := -\frac{\mu_2^2}{12}. \end{array} \right.$$

De plus, définissons les cellules suivantes :

$$\left\{ \begin{array}{l} C_1 := \{\mu_2 \leq 0, \quad \mu_1 \leq -\mu_1^*, \quad \mu_0 \leq r_1\}, \\ C_2 := \{\mu_2 \leq 0, \quad \mu_1 \in [-\mu_1^*; 0[, \quad \mu_0 \in ]-\infty; r_1] \setminus \{\mu_0^*\}\}, \\ C_3 := \{\mu_2 \leq 0, \quad \mu_1 \in ]0; \mu_1^*], \quad \mu_0 \in ]\mu_0^*; r_2]\}, \\ C_4 := \{\mu_2 \geq 0, \quad \mu_1 < 0, \quad \mu_0 \leq r_1\}, \\ C_5 := \{\mu_2 \leq 0, \quad \mu_1 \in ]0; \mu_1^*], \quad \mu_0 \in ]-\infty; \mu_0^* \cup [r_2; r_1]\}, \\ C_6 := \{\mu_2 \leq 0, \quad \mu_1 \geq \mu_1^*, \quad \mu_0 \leq r_1\}, \\ C_7 := \{\mu_2 \geq 0, \quad \mu_1 > 0, \quad \mu_0 \leq r_1\}, \\ C_8 := \{\mu_2 \leq 0, \quad \mu_1 \in [-\mu_1^*; 0[), \\ C_9 := \{\mu_2 \geq 0, \quad \mu_1 > 0\}. \end{array} \right.$$

Nous distinguons alors les cas suivants.

- Supposons  $\mu_1 = 0$ . Alors  $b_3(1, 1)$  est la racine réelle maximale de  $\tilde{p}$ .
- Supposons  $\mu_1 \neq 0$  et  $\mu_0 \neq \mu_0^*$ . Alors  $b_3(1, 1)$  est la racine réelle maximale de  $\tilde{p}$  si et seulement si

$$\mu \in C_1 \cup C_2 \cup C_3 \cup C_4,$$

et  $b_3(1, -1)$  est racine réelle maximale de  $\tilde{p}$  si et seulement si :

$$\mu \in C_5 \cup C_6 \cup C_7.$$

- Supposons  $\mu_1 \neq 0$  et  $\mu_0 = \mu_0^*$ . Alors,  $b_3(1, -1)$  est racine réelle maximale de  $\tilde{p}$  si et seulement si

$$\mu \in C_8 \cup C_9$$

et  $b_3(1, 1)$  est racine réelle maximale de  $\tilde{p}$  dans le cas contraire.

Ces résultats sont résumés dans la Table 3.7.

*Démonstration.* • Soit  $n = 2$ . Trivialement, d'après la Table 3.5, la racine réelle maximale est donnée par  $b_1(1)$  étant donné que l'intérieur de la racine carrée est positif.

- Soit  $n = 3$ . Nous procédons en deux étapes.

- Supposons  $\mu_2^2 + 12\mu_0 \neq 0$ . Le calcul de CellDecomposition pour  $\tilde{p}(b_2, b_0, a, c)$  avec la restriction  $b_0 > 0$  et  $\mu_2^2 + 12\mu_0 \neq 0$  et ordre de projection

$$c_2 \rightarrow c_1 \rightarrow c_0 \rightarrow a_2 \rightarrow a_1 \rightarrow a_0 \rightarrow b_0,$$

permet de décomposer l'espace des paramètres en 28 cellules dans lesquelles le nombre de racines réelles de  $\tilde{p}$  est constant et dont l'expression par radicaux de la racine réelle maximale de  $\tilde{p}$  ne change pas. Pratiquement, la commande *Maple* utilisée est :

$$\text{CellDecomposition}([\tilde{p} = 0, b_0 > 0, \mu_2^2 + 12\mu_0 \neq 0], [b_2], [b_0, a_0, a_1, a_2, c_0, c_1, c_2]).$$

L'évaluation des racines par radicaux données en Table 3.5 montre que, dans chaque cellule, soit  $b_2(1, 1)$  est maximale, soit il n'y a pas de solution. Étant donné que nous avons considéré  $b_0$  comme un paramètre sans restriction ( $b_0$  vérifie  $b_0^2 = a_0^2 + c_0^2$ ), il est normal de voir apparaître des cellules sans solution. Celles-ci ne correspondent alors à aucune situation réelle, puisque nous sommes assurés que  $\mathcal{R} = 0$  possède une solution. Nous pouvons alors éliminer ces cellules et ne considérer que celles disposant au moins d'une solution réelle. Ainsi,  $b_2(1, 1)$  est racine réelle maximale de  $\tilde{p}$  si  $\mu_2^2 + 12\mu_0 \neq 0$ .

- Supposons  $\mu_2^2 + 12\mu_0 = 0$ . La réduction de  $\tilde{p}$  modulo  $\mu_2^2 + 12\mu_0$  en éliminant  $c_0$  donne :

$$\tilde{p}^* := 3b_2^4 + 6(2a_1 - a_2^2 - c_2^2)b_2^2 - 24b_0b_2 - 2(-2a_2^2 - 2c_2^2 + 4a_1)^2.$$

Remarquons alors que le polynôme obtenu ne dépend plus que des paramètres  $a_1, a_2, c_2$  et  $b_0$ . En utilisant la commande *Maple*

$$\text{CellDecomposition}([\tilde{p}^* = 0, b_0 > 0], [b_2], [a_1, a_2, b_0, c_2]),$$

nous décomposons l'espace des paramètres  $(a, c, b_0)$  restreint à  $\mu_2^2 + 12\mu_0 = 0$  en 2 cellules dans lesquelles le nombre de racines réelles de  $\tilde{p}^*$  est constant. L'évaluation des racines par radicaux montre alors que  $b_2(1, 1)$  est racine réelle maximale dans chaque cellule.

$b_2(1, 1)$  est alors la racine réelle maximale de  $\tilde{p}$  quelles que soient les valeurs des paramètres  $(a, c, b_0)$  avec  $b_0 > 0$ .

- Soit  $n = 4$ .  $\tilde{p}$  est un polynôme univarié de degré 4 en  $b_3^2$ . Notons  $t := b_3^2 - d_6$  et le polynôme  $\tilde{p}_t$

$$\tilde{p}_t := t^4 + \mu_2 t^2 + \mu_1 t + \mu_0,$$

où les coefficients  $\mu_i$  sont définis à la Table 3.5. Notons que si  $t_{\max}$  est racine réelle maximale de  $\tilde{p}_t$ , alors  $\sigma := \sqrt{t_{\max} + d_6}$  est racine réelle maximale de  $\tilde{p}$ .

Par ailleurs, décomposer l'espace des paramètres en cellules via la variété discriminante de  $\tilde{p}_t$  en les paramètres  $(a, c, b_0)$  est très couteux à cause du grand nombre de paramètres, et n'est pas réalisable sur nos machines. Nous proposons alors de décomposer l'espace des paramètres selon des valeurs génériques de  $\mu_i, i = 0, 1, 2$ , définis dans la Table 3.5. Une décomposition en cellules de l'espace des paramètres  $(\mu_0, \mu_1, \mu_2)$  pour un polynôme générique de degré 4 est donné en Annexe D et est reprise ici afin de terminer la preuve pour  $n = 4$ .  $\square$

**Remarque 3.29.** Notons que le cas  $n = 1$  est un cas particulier et la Proposition 3.11 permet de conclure immédiatement quant à la solution définie positive de l'équation de Riccati (3.11).

De telles solutions sont facilement exploitables, notamment par un calculateur embarqué, ce qui motive la présente démarche. Une limitation évidente de l'approche développée est l'ordre du système  $G$ . Cependant, pour certains systèmes remarquables, nous pouvons augmenter l'ordre du système comme le montre la section suivante.

$n$	Matrice $X$ solution de $\mathcal{R} = 0$
1	$b_0 - a_0$
2	$\begin{pmatrix} b_0 b_1 - (a_0 a_1 + c_0 c_1) & b_0 - a_0 \\ b_0 - a_0 & b_1 - a_1 \end{pmatrix}$
3	$\begin{pmatrix} b_0 b_1 - (a_0 a_1 + c_0 c_1) & b_0 b_1 - (a_0 a_2 + c_0 c_2) & b_0 - a_0 \\ b_0 b_1 - (a_0 a_2 + c_0 c_2) & b_1 b_2 - (a_1 a_2 - c_1 c_2) - (b_0 - a_0) & b_1 - a_1 \\ b_0 - a_0 & b_1 - a_1 & b_2 - a_2 \end{pmatrix}$
4	$\begin{pmatrix} b_0 b_1 & b_0 b_2 & b_0 b_3 & b_0 - a_0 \\ -(a_0 a_1 + c_0 c_1) & -(a_0 a_2 + c_0 c_2) & -(a_0 a_3 + c_0 c_3) & \\ b_0 b_2 & b_1 b_2 - b_0 b_3 + a_0 a_3 & b_1 b_3 - b_0 & b_1 - a_1 \\ -(a_0 a_2 + c_0 c_2) & +c_0 c_3 - (a_1 a_2 + c_1 c_2) & +a_0 - (a_1 a_3 + c_1 c_3) & \\ b_0 b_3 & b_1 b_3 - b_0 + a_0 & b_1 b_2 - b_2 b_3 - b_1 & b_2 - a_2 \\ -(a_0 a_3 + c_0 c_3) & -(a_1 a_3 + c_1 c_3) & +a_1 - (a_2 a_3 + c_2 c_3) & \\ b_0 - a_0 & b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \end{pmatrix}$

 TABLE 3.1 – Matrices  $X$  solution de  $\mathcal{R} = 0$  pour  $1 \leq n \leq 4$ 

$n$	Système polynomial $\mathcal{B}$
1	$\mathcal{B}_0 := b_0^2 - d_0 = 0$
2	$\mathcal{B}_0 := b_0^2 - d_0 = 0$ $\mathcal{B}_1 := b_1^2 - 2b_0 - d_2 = 0$
3	$\mathcal{B}_0 := b_0^2 - d_0 = 0$ $\mathcal{B}_1 := b_1^2 - 2b_0 b_2 - d_2 = 0$ $\mathcal{B}_2 := b_2^2 - 2b_1 - d_4 = 0$
4	$\mathcal{B}_0 := b_0^2 - d_0 = 0$ $\mathcal{B}_1 := b_1^2 - 2b_0 b_2 - d_2 = 0$ $\mathcal{B}_2 := b_2^2 - 2b_1 b_3 + 2b_0 - d_4 = 0$ $\mathcal{B}_3 := b_3^2 - 2b_2 - d_6 = 0$

 TABLE 3.2 – Système polynomial  $\mathcal{B}$  pour  $1 \leq n \leq 4$

$n$	Variables $d_{2k}$
1	$d_0 := a_0^2 + c_0^2$
2	$d_0 := a_0^2 + c_0^2$ $d_2 := (a_1^2 + c_1^2) - 2a_0$
3	$d_0 := a_0^2 + c_0^2$ $d_2 := (a_1^2 + c_1^2) - 2(a_0 a_2 + c_0 c_2)$ $d_4 := (a_2^2 + c_2^2) - 2a_1$
4	$d_0 := a_0^2 + c_0^2$ $d_2 := a_1^2 + c_1^2 - 2(a_0 a_2 + c_0 c_2)$ $d_4 := a_2^2 + c_2^2 - 2(a_1 a_3 + c_1 c_3) + 2a_0$ $d_6 := a_3^2 + c_3^2 - 2a_2$

 TABLE 3.3 – Expression des  $d_{2k}$  en fonction de  $a$  et  $c$  pour  $1 \leq n \leq 4$ 

$n$	$\tilde{\mathcal{R}}$
1	$\emptyset$
2	$b_1 = b_1$ $\tilde{p} := b_1^2 - 2b_0 - d_2 = 0$
3	$b_1 = \frac{1}{2}(b_2^2 - d_4)$ $b_2 = b_2$ $\tilde{p} := b_2^4 - 2d_4 b_2^2 - 8b_0 b_2 + d_4^2 - 4d_2 = 0$
4	$b_1 = \begin{cases} \pm\sqrt{d_2 - b_0 d_6} & \text{si } b_3 = 0 \\ \frac{b_3^4 - 2d_6 b_3^2 + 8b_0 + d_6^2 - 4d_4}{8b_3} & \text{sinon} \end{cases}$ $b_2 = \frac{1}{2}(b_3^2 - d_6)$ $b_3 = b_3$ $\tilde{p} = b_3^8 - 4d_6 b_3^6 + (6d_6^2 - 8d_4 - 48b_0)b_3^4 + 4(-d_6^3 + 4(d_4 + 2b_0)d_6 - 16d_2)b_3^2 + (d_6^2 - 4d_4 + 8b_0)^2$

 TABLE 3.4 – Paramétrisation rationnelle  $\tilde{\mathcal{R}}$  des solutions de  $\tilde{\mathcal{E}}$  pour  $1 \leq n \leq 4$

$n$	$b_{n-1}(b_0, d_{2k})$
1	$\emptyset$
2	$\varepsilon := \pm 1$ $b_1(\varepsilon) = \varepsilon \sqrt{2(b_0 - a_0) + a_1^2 + c_1^2}$
3	$\varepsilon_1 := \pm 1 \quad \varepsilon_2 := \pm 1 \quad \varepsilon := (\varepsilon_1, \varepsilon_2)$ $\mu_0 := d_4^2 - 4d_2 \quad \mu_1 := -8b_0 \quad \mu_2 := -2d_4$ $p_4 := 4d_2 - \frac{4}{3}d_4^2 \quad q_4 := \frac{8}{3}d_2d_4 - \frac{16}{27}d_4^3 - 8b_0^2$ $\delta_4 := 27(4p_4^3 + 27q_4^2) \quad \alpha_4 := \left(\frac{1}{2}(\sqrt{\delta_4} - 27q_4)\right)^{(1/3)}$ $u_4 := \begin{cases} \frac{1}{3}\left(\alpha_4 - \frac{3p_4}{\alpha_4} - \mu_2\right) & \text{si } p_4 \neq 0 \\ (-q_4)^{1/3} - \frac{\mu_2}{3} & \text{si } p_4 = 0 \end{cases}$ $\Delta_4 := -\left(u_4 + \mu_2 + \varepsilon_1 \frac{\mu_1}{\sqrt{2u_4}}\right)$ $b_2(\varepsilon) = \frac{\sqrt{2}}{2} (\varepsilon_1 \sqrt{u_4} + \varepsilon_2 \sqrt{\Delta_4})$
4	$\varepsilon_1 := \pm 1 \quad \varepsilon_2 := \pm 1 \quad \varepsilon_3 := \pm 1 \quad \varepsilon := (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ $\mu_0 := 16(2b_0 - d_4)^2 - 64d_2d_6 \quad \mu_1 := -64(d_2 + b_0d_6) \quad \mu_2 := -8(6b_0 + d_4)$ $p_4 := -\left(\frac{\mu_2^2}{12} + \mu_0\right) \quad q_4 := \frac{\mu_2}{3}\left(\mu_0 - \frac{\mu_2^2}{36}\right) - \frac{\mu_1^2}{8}$ $\delta_4 := 27(4p_4^3 + 27q_4^2) \quad \alpha_4 := \left(\frac{1}{2}(\sqrt{\delta_4} - 27q_4)\right)^{(1/3)}$ $u_4 := \begin{cases} \frac{1}{3}\left(\alpha_4 - \frac{3p_4}{\alpha_4} - \mu_2\right) & \text{si } p_4 \neq 0 \\ (-q_4)^{1/3} - \frac{\mu_2}{3} & \text{si } p_4 = 0 \end{cases}$ $\Delta_4 := -\left(u_4 + \mu_2 + \varepsilon_1 \frac{\mu_1}{\sqrt{2u_4}}\right)$ $t(\varepsilon_1, \varepsilon_2) := \begin{cases} \frac{\sqrt{2}}{2} (\varepsilon_1 \sqrt{u_2} + \varepsilon_2 \sqrt{\Delta_2}) & \text{si } \mu_1 \neq 0 \\ \varepsilon_1 \frac{\sqrt{2}}{2} \sqrt{\varepsilon_2 \sqrt{p^2 - 4r} - p} & \text{si } \mu_1 = 0 \end{cases}$ $b_3(\varepsilon) = \varepsilon_3 \sqrt{t(\varepsilon_1, \varepsilon_2) + d_6}$

 TABLE 3.5 – Racines par radicaux de  $\tilde{p}$  pour  $1 \leq n \leq 4$

$n$	$\mathcal{V}_s$
1	$\emptyset$
2	$\emptyset$
3	$\{(a, c) \in \mathbb{R}^{2n} \mid p_4(a, c) = 0\}$
4	$\{(a, c) \in \mathbb{R}^{2n} \mid p_4(a, c) = 0\} \cup \{(a, c) \in \mathbb{R}^{2n} \mid \mu_1(a, c) = 0\}$

TABLE 3.6 – Variétés singulières introduites par les racines par radicaux de  $\tilde{p}$

$n$	Racine réelle maximale $\tilde{\sigma}$ de $\tilde{p}(b_{n-1})$
1	$\emptyset$
2	$\tilde{\sigma} := b_1(1) = \sqrt{2b_0 + a_1^2 + c_1^2 - 2a_0}$
3	$\tilde{\sigma} := b_2(1, 1)$
4	$\tilde{\sigma} := \begin{cases} b_3(1, 1, 1) & \text{si } \mu_1 = 0 \text{ ou } \mu \in \{C_1 \cup C_2 \cup C_3 \cup C_4\} \text{ ou } \mu \notin \{C_8 \cup C_9\}, \\ b_3(1, -1, 1) & \text{si } \mu \in \{C_5 \cup C_6 \cup C_7\} \text{ ou } \mu \in \{C_8 \cup C_9\}. \end{cases}$

TABLE 3.7 – Expression par radicaux de la racine réelle maximale de  $\tilde{p}$  pour  $2 \leq n \leq 4$

## 3.7 Cas particulier des systèmes comportant une approximation de Padé

Avant de conclure ce chapitre, nous proposons d'étendre l'approche aux systèmes à retard. Il existe une vaste littérature sur la commande robuste des systèmes à retard (voir, par exemple, [FÖT96, PG90]). Dans cette section, nous proposons d'approximer les retards par des fractions rationnelles appelées *approximations de Padé*. Dans ce contexte des systèmes linéaires à dimension finie, nous montrons que la factorisation spectrale permettant la résolution de l'ERA possède des propriétés intéressantes simplifiant sa résolution.

### 3.7.1 Rappels sur les approximations de Padé

Un *approximant de Padé* de la fonction exponentielle est une fraction rationnelle  $N_{p,q}/D_{p,q}$ , où  $N_{p,q}$  désigne un polynôme de degré  $p$  et  $D_{p,q}$  de degré  $q$ , telle que le développement limité de la fraction à l'ordre  $p+q$  soit identique à celui de l'exponentielle [Pad92]. Plus exactement, cette fraction est définie comme suit.

**Théorème 3.20.** [BGM96, Pad92, Mar77, Lam93] Soit  $(p, q) \in \mathbb{N}^{*2}$ . Soient  $n_{p,q} \in \mathbb{Q}[s]$  et  $d_{p,q} \in \mathbb{Q}[s]$  tels que :

$$N_{p,q}(s) := \sum_{i=0}^p \frac{p!}{(p-i)!} \frac{(p+q-i)!}{(p+q)!} \frac{s^i}{i!},$$

$$D_{p,q}(s) := \sum_{j=0}^q (-1)^j \frac{q!}{(q-j)!} \frac{(p+q-j)!}{(p+q)!} \frac{s^j}{j!}.$$

Alors,  $e^s$  s'approxime par

$$e^s \simeq \frac{n_{p,q}(s)}{d_{p,q}(s)},$$

que l'on appelle *approximant de Padé*.

Un approximant de Padé vérifie la propriété suivante.

**Théorème 3.21.** [Mar77] L'approximation de Padé d'ordre  $(p, q)$  de  $e^{-s}$  est stable pour  $q \geq p - 2$ , c'est-à-dire ne contient pas de pôles dans le demi-plan ouvert droit  $\mathbb{C}_+$ . De plus, nous avons :

$$N_{p,p}(s) = D_{p,p}(-s).$$

En particulier, nous utiliserons principalement les approximations de Padé dites "équilibrées", c'est-à-dire de même degré en  $s$  au numérateur et au dénominateur. En effet, celles-ci permettent d'approximer en basse fréquence le déphasage dû à un retard, sans modification du gain. Nous avons alors le corollaire suivant du théorème précédent.

**Corollaire 3.2.** Une approximation de Padé d'ordre  $(p, p)$  est stable, c'est-à-dire que ses pôles sont tous à partie réelle strictement négative. Nous parlerons alors d'approximation de Padé *équilibrée* d'ordre  $p$ .

Autrement dit, si  $G_p = N_p/D_p$  désigne une approximation de Padé équilibrée d'ordre  $p$ , alors :

$$N_p(s) = D_p(-s), \quad \{s_i \in \mathbb{C} \mid D_p(s_i) = 0\} \subset \mathbb{C}_-, \quad \{s_i \in \mathbb{C} \mid N_p(s_i) = 0\} \subset \mathbb{C}_+.$$

### 3.7.2 Approximations de Padé et factorisation spectrale

Soit  $G_p$  une approximation de Padé équilibrée d'un retard  $e^{-\tau s}$  définie par :

$$G_p(s) := \frac{N_p(s)}{D_p(s)}.$$

Rappelons que nous avons  $N_p(s) = D_p(-s)$ . Soit un système  $G$  linéaire, SISO, LTI, de la forme suivante

$$G(s) = G_0(s) G_p(s),$$

où  $G_0 := N_0/D_0$  est une fonction de transfert strictement propre et où  $G$  ne contient pas de simplification pôle/zéro. Notons alors :

$$\begin{cases} N(s) & := N_0(s) N_p(s), \\ D(s) & := D_0(s) D_p(s). \end{cases}$$

Il en résulte que  $G = N/D$  avec  $\text{gcd}(N, D) = 1$ , où  $\text{gcd}(N, D)$  désigne le plus grand commun diviseur de  $N$  et  $D$ . Notons  $n$  l'ordre de  $G$ .

Par ailleurs, à la Remarque 3.14, nous avons vu que calculer les valeurs propres de la matrice Hamiltonienne  $\mathcal{H}$  revient à calculer les racines du polynôme pair suivant :

$$f(s) = (-1)^n (N(s) N(-s) + D(s) D(-s)).$$

Dans le cas présent, nous avons :

$$f(s) = (-1)^n (N_0(s) N_p(s) N_0(-s) N_p(-s) + D_0(s) D_p(s) D_0(-s) D_p(-s)).$$

Or, nous avons  $N_p(s) = D_p(-s)$ , ce qui donne :

$$f(s) = (-1)^n D_p(-s) D_p(s) (N_0(s) N_0(-s) + D_0(s) D_0(-s)).$$

Or, d'après le Corollaire 3.2, nous savons que  $D_p$  a toutes ses racines dans le demi-plan ouvert gauche  $\mathbb{C}_-$ . Ainsi, trouver une factorisation spectrale de  $f$  revient à trouver une factorisation spectrale du polynôme pair  $f_0$  défini par

$$f_0(s) := (-1)^{(n-p)} (N_0(s) N_0(-s) + D_0(s) D_0(-s)),$$

$f_0$  étant uniquement défini par la partie sans retard de  $G$ . En effet, soit  $g_0$  tel que :

$$f_0(s) = (-1)^{(n-p)} g_0(s) g_0(-s).$$

Alors, nous avons :

$$f(s) = (-1)^p D_p(s) D_p(-s) (-1)^{(n-p)} g_0(s) g_0(-s).$$

Une factorisation spectrale de  $f$  est alors donnée par

$$f(s) = (-1)^n g(s) g(-s),$$

où  $g(s) := g_0(s) D_p(s)$ . Nous récapitulons ce résultat dans le théorème suivant.

**Théorème 3.22.** Soit  $G = G_0 \frac{N_p}{D_p}$  un système linéaire continu, SISO, LTI, où  $N_p(s) = D_p(-s)$ . Alors, le problème de factorisation spectrale associé à  $G$  se réduit au problème de factorisation spectrale de la fonction de transfert  $G_0$ .

## 3.8 Conclusion

Dans ce chapitre, nous avons proposé deux méthodes permettant de résoudre une équation de Riccati algébrique  $\mathcal{R} = 0$  à paramètres indéterminés.

La première propose de passer par la matrice Hamiltonienne associée, dont les sous-espaces propres possèdent des propriétés intéressantes exploitables par factorisation spectrale. Ainsi, par une paramétrisation des vecteurs propres de cette matrice Hamiltonienne, et en considérant une factorisation spectrale de son polynôme caractéristique, nous avons montré comment paramétrer les solutions de l'équation de Riccati algébrique suivant  $n$  variables issues d'un système polynomial  $\mathcal{E}$ .

La seconde méthode propose d'étudier directement l'équation de Riccati comme un système polynomial à  $n(n+1)/2$  variables. Par un schéma de substitutions, nous montrons que ce système polynomial se réduit à l'étude du même système polynomial  $\mathcal{E}$  obtenu par la première méthode.

Ces deux méthodes sont complémentaires. La première nous permet de déduire un critère algébrique permettant d'identifier la solution définie positive, en étudiant les propriétés de la factorisation spectrale. La seconde propose des expressions explicites des composantes de la matrice  $X$  suivant les variables du système  $\mathcal{E}$ . Ce dernier aspect permet d'éviter de nombreux calculs inhérents à la première méthode. Pour n'importe quel système SISO d'ordre  $n$  et sans aucun calcul, nous obtenons alors une paramétrisation des solutions de  $\mathcal{R} = 0$  suivant les solutions d'un système polynomial  $\mathcal{E}$ .

L'étude de l'équation de Riccati algébrique étant réduite au système  $\mathcal{E}$ , nous avons montré qu'il était préférable d'utiliser une Représentation Univariée Rationnelle de celui-ci pour paramétrer ses solutions. Le calcul d'une RUR augmente avec l'ordre du système et le nombre de paramètres, comme nous le montrons en Annexe E.

Disposant d'une paramétrisation des solutions de l'équation de Riccati suivant les racines d'un polynôme univarié  $p$ , il suffit d'isoler la racine réelle maximale de celui-ci afin d'obtenir la solution définie positive de l'équation de Riccati. Pour des équations de taille  $n$  inférieure à 4, nous avons alors montré que cette solution pouvait être trouvée sous forme close en utilisant des formules par radicaux et une décomposition de l'espace des paramètres à l'aide de la variété discriminante de  $p$ .

Enfin, nous avons aussi montré qu'en considérant des systèmes contenant des approximations de Padé, le problème de factorisation spectrale se simplifie en ne laissant apparaître que la partie de la factorisation correspond à la partie sans approximant de Padé. Ainsi, la résolution de l'équation de Riccati pour des systèmes présentant une approximation de Padé est grandement simplifiée.

L'ensemble du processus nous donne une méthode en théorie valable pour n'importe quel système  $G$  d'ordre  $n$  à paramètres indéterminés. Cependant, en pratique, les différentes paramétrisations sont coûteuses à calculer et l'ordre du système est très vite limitant. De plus, lorsque nous cherchons des formules explicites pour les solutions de  $\mathcal{R} = 0$ , nous sommes fortement contraints par l'ordre du système  $G$ . Ces limitations restent les principales difficultés de cette méthode.

En revanche, lorsque ces paramétrisations sont trouvées pour des paramètres indéterminés, le problème est considéré comme résolu et peut être exploité à souhait. Nous voyons alors de potentielles applications en commande LQG, le problème étant alors complètement résolu pour des systèmes d'ordre faible à l'aide de notre approche. De plus, pour des systèmes comportant des approximations de Padé, cette approche reste applicable, ce qui élargit le champ de systèmes pouvant être étudiés.

Le Chapitre 2 montre que l'équation de Riccati intervient dans le problème de commande  $H_\infty$ . Le Chapitre 4 montre comment nous pouvons étendre les résultats sur les équations de Riccati algébriques à paramètres indéterminés présentés dans ce chapitre à l'étude de la commande  $H_\infty$ .



# Chapitre 4

## Approche symbolique de la commande $H_\infty$ par Loop-Shaping

Le chapitre précédent a montré comment réduire l'étude de l'équation de Riccati  $\mathcal{R} = 0$  à un système polynomial en les variables  $\{b_0, \dots, b_{n-1}\}$ , puis comment trouver une paramétrisation rationnelle de ce dernier système en fonction de  $b_{n-1}$ . Nous avons ensuite caractérisé de manière explicite la solution définie positive de l'ERA à l'aide de  $b_{n-1}$ .

L'équation de Riccati ne constitue qu'une première étape à la résolution du problème de commande robuste  $H_\infty$  (PCR) introduit à la Section 2.3.2 du Chapitre 2. Il reste maintenant à calculer les valeurs propres d'une matrice construite à partir de la solution de l'équation de Riccati algébrique précédente. Ce calcul de valeurs propres est réalisé à l'aide du calcul d'un polynôme caractéristique.

Pour des systèmes d'ordre  $n$  avec  $n$  inférieur ou égal à 4, nous disposons d'une forme close de la solution définie positive de l'équation de Riccati algébrique issue du PCR. Pour de tels  $n$ , le calcul de polynôme caractéristique précédent peut alors être explicité pour les systèmes les plus généraux (c'est-à-dire lorsque tous les coefficients de la fonction de transfert sont considérés comme des paramètres indéterminés). Nous pouvons alors caractériser les racines par radicaux de ces polynômes caractéristiques et identifier les racines maximales à l'aide du concept de variété discriminante (voir Section 3.1.10). Disposant de ces racines maximales sous forme close, nous pouvons déduire une solution du PCR sous forme close, ainsi que les correcteurs optimaux associés. En revanche, pour  $n \geq 5$ , une étape numérique est nécessaire du fait de la théorie de Galois. En effet, celle-ci stipule que toute racine d'un polynôme quelconque de degré 5 ou plus n'est généralement pas exprimable par radicaux (voir, par exemple, [Mut80, Tig15]). Dans ce chapitre, nous détaillons les deux approches.

Nous présentons ensuite une méthode pour régler une pondération scalaire Loop-Shaping de manière optimale pour deux exemples : un système masse-ressort sans et avec amortissement. Celle-ci est basée sur le calcul d'un polynôme encodant toutes les valeurs propres des matrices, non nécessairement définies positives, solutions des équations de Riccati algébriques intervenant dans le problème  $H_\infty$ .

Les résultats de ce chapitre sont publiés dans les articles de conférence [RBQQ17] et [RBQ<sup>+</sup>17].

## Sommaire

<b>4.1</b>	<b>Résolution explicite du Problème de Commande Robuste . . .</b>	<b>128</b>
4.1.1	Calcul explicite du critère $\gamma_{\text{opt}}$ . . . . .	128
4.1.2	Correcteurs $H_\infty$ paramétriques . . . . .	138
4.1.3	Exemples . . . . .	140
<b>4.2</b>	<b>Résolution symbolique/numérique du problème de commande robuste . . . . .</b>	<b>152</b>
<b>4.3</b>	<b>Réglage optimal d'une pondération Loop-Shaping . . . . .</b>	<b>155</b>
4.3.1	Une méthode de réglage de pondération scalaire de type gain . . .	156
4.3.2	Application au système double masse-ressort . . . . .	159
<b>4.4</b>	<b>Commande <math>H_\infty</math> adaptative indirecte . . . . .</b>	<b>170</b>
4.4.1	Système double masse-ressort sans amortissement bouclé avec correction adaptative globale . . . . .	171
4.4.2	Système double masse-ressort avec amortissement bouclé avec correction adaptative locale . . . . .	171
<b>4.5</b>	<b>Conclusion . . . . .</b>	<b>173</b>

---

## 4.1 Résolution explicite du Problème de Commande Robuste

La Section 3.6.4 du chapitre précédent donne des formes closes de la solution  $X > 0$  de l'équation de Riccati algébrique  $\mathcal{R} = 0$  associée à un système  $G$  d'ordre  $n$ , avec  $1 \leq n \leq 4$ . Les formes closes  $X$  obtenues sont des expressions explicites dépendant des paramètres  $a$  et  $c$  de  $G$ . Disposant de ces formes closes, cette section montre comment expliciter des formules par radicaux de la solution du PCR (voir Définition 2.7). Dans toute cette section, nous supposons alors  $1 \leq n \leq 4$ .

### 4.1.1 Calcul explicite du critère $\gamma_{\text{opt}}$

Afin de déterminer la valeur du critère  $H_\infty$ , noté  $\gamma_{\text{opt}}$ , il reste à déterminer les valeurs propres d'une matrice, construite à l'aide de  $X > 0$  et  $\mathcal{Q}$  (voir Théorème 2.6). A paramètres  $a$  et  $c$  fixés, rappelons que celui-ci peut se déterminer de deux manières différentes (voir (2.19))

$$\gamma_{\text{opt}} = \sqrt{1 + \lambda_{\max}(YX)} = \sqrt{1 + \lambda_\star^2(\mathcal{Q}X)},$$

où  $\mathcal{Q}$  est une matrice explicitement connue dépendant des paramètres  $a$  et  $c$  (voir le Théorème 2.6),  $Y := \mathcal{Q}X\mathcal{Q}$ ,  $\lambda_\star(M)$  désigne la valeur absolue maximale des valeurs propres réelles de  $M$  et  $\lambda_{\max}(M)$  la valeur propre réelle maximale de  $M$ . De plus, rappelons que les matrices  $YX$  et  $\mathcal{Q}X$  vérifient les propriétés suivantes (voir l'Annexe C).

**Proposition 4.1.** Les valeurs propres de  $YX$  sont réelles strictement positives et les valeurs propres de  $\mathcal{Q}X$  sont réelles.

*Démonstration.* Voir preuve du Théorème 2.6 et la Section C.2 de l'Annexe C. □

A paramètres fixés, le calcul de  $\gamma_{\text{opt}}$  se ramène donc à un calcul de valeurs propres de  $Y X$  ou de  $\mathcal{Q} X$ . Supposons maintenant que les paramètres  $a$  et  $c$  ne sont pas fixés à des valeurs numériques et définissons alors les polynômes caractéristiques suivants :

$$h := \det(Y X - \lambda I), \quad (4.1)$$

$$h_\star := \det(\mathcal{Q} X - \lambda I). \quad (4.2)$$

Notons que  $h$  et  $h_\star$  sont deux polynômes en  $\lambda$ , dont les coefficients sont des expressions par radicaux en  $a$  et  $c$ . Afin de déterminer  $\gamma_{\text{opt}}$  de manière explicite en fonction de  $a$  et  $c$ , nous disposons alors de deux méthodes.

La première consiste à étudier les racines de  $h_\star$ . Comme  $n \leq 4$ , la matrice  $\mathcal{Q} X$  est au plus de taille  $4 \times 4$  et donc  $h_\star$  est de degré au plus 4 en  $\lambda$ . Nous pouvons alors expliciter les racines de  $h_\star$  par radicaux (voir l'Annexe D). Cependant, le travail de discrimination de la racine de plus grande valeur absolue de  $h_\star$  est complexe car les coefficients de  $h_\star$  (c'est-à-dire les paramètres de  $G$ ) sont inconnus.

La deuxième consiste à étudier les racines de  $h$ . Là encore, comme  $n \leq 4$ , la matrice  $Y X$  est au plus de taille  $4 \times 4$  et donc  $h$  est de degré au plus 4 en  $\lambda$ , et nous pouvons expliciter ses racines par radicaux (voir Annexe D). Ici, le travail de discrimination est plus facile puisque nous cherchons la racine réelle maximale sur l'espace des paramètres. Notamment, ce travail s'effectue à l'aide d'un calcul de variétés discriminantes comme nous l'avons vu au chapitre précédent. Rappelons que, pour un polynôme à coefficients génériques de degré inférieur ou égal à 4, l'Annexe D donne l'expression par radicaux de la racine réelle maximale de ce polynôme suivant les valeurs de ses coefficients. En particulier, cette annexe fournit l'expression par radicaux de la racine réelle maximale d'un polynôme à coefficients génériques dont les racines sont toutes réelles (voir Proposition D.1, Corollaire D.1 et Corollaire D.2 de l'Annexe D).

Dans cette section, nous privilégierons donc cette deuxième méthode, qui s'avère plus simple que de chercher le maximum des valeurs absolues. Par conséquent, reformulons le PCR à paramètres non fixés.

**Définition 4.1. Étude du critère de robustesse à paramètres non fixés**

Trouver une expression sous forme close  $\lambda_m$  de la racine réelle maximale de  $h$ , défini par (4.1), dépendant de  $a$  et  $c$  afin d'en déduire  $\gamma_{\text{opt}} := \sqrt{1 + \lambda_m}$ .

**Remarque 4.1.** Dans toute cette section,  $X$ ,  $\mathcal{Q}$ ,  $Y$ ,  $\lambda_m$ ,  $\gamma_{\text{opt}}$ , sont des fonctions des paramètres  $a$  et  $c$ . Afin d'alléger l'écriture, nous ne précisons pas leur dépendance en ces paramètres lorsque cela n'est pas nécessaire.

**Remarque 4.2.** Notons que pour  $n \geq 5$ , les matrices  $X$  et  $\mathcal{Q}$  sont de taille  $n \times n$  supérieure ou égal à  $5 \times 5$ . Ainsi, les polynômes  $h$  et  $h_\star$  sont de degré supérieur ou égal à 5 en  $\lambda$ . S'ils ne sont pas factorisables, ils sont alors de degré trop élevé pour que l'on puisse trouver des formes closes par radicaux de leurs racines. En effet, la théorie de Galois (voir, par exemple, [Mut80]) montre qu'au delà du degré 4, il est généralement impossible d'exprimer les solutions de polynômes par radicaux. Dans ce cas, il est alors impossible d'en déduire une expression sous forme close par radicaux de  $\gamma_{\text{opt}}$  et d'un correcteur  $H_\infty$  satisfaisant ce critère en fonction de  $a$  et  $c$ . Durant cette thèse, nous avons pu néanmoins observer grâce à de nombreux tests sur des exemples intéressants en pratique que certains systèmes présentant des "symétries"

pouvaient voir leurs polynômes associés  $\tilde{p}$  (voir Section 3.6.2) et  $h$  se factoriser. Ainsi, nous invitons le lecteur voulant appliquer nos méthodes à bien regarder les polynômes pour les systèmes considérés avant de conclure à l'absence de formes closes des solutions par radicaux.

Cherchons donc à expliciter les racines de  $YX$ . Rappelons que la condition de non observabilité est  $\Delta_o = 0$ . Celle-ci est explicitée dans la Table 4.1 et est utilisée pour calculer  $\mathcal{Q}$  (voir (2.13) et le Théorème 2.5 de la Section 2.3.4). Disposant de  $X > 0$  et de  $\mathcal{Q}$ , nous en déduisons  $Y := \mathcal{Q}X\mathcal{Q}$ . Dans la Table 4.2 nous donnons des formules explicites pour ces mêmes matrices  $\mathcal{Q}$  et, dans la Table 4.3, les formules explicites pour  $Y$ . Enfin, la Table 4.4 liste les polynômes  $h$ . Le théorème suivant donne alors une forme close de  $\gamma_{\text{opt}}$  pour  $1 \leq n \leq 4$ .

**Théorème 4.1.** Pour  $1 \leq n \leq 4$ ,  $\gamma_{\text{opt}}$  est explicitable par radicaux en fonction des paramètres. Les expressions sous forme close de  $\gamma_{\text{opt}}$  sont données dans la Table 4.5.

*Démonstration.* Tout d'abord, comme nous le rappelons à la Proposition 4.1 ci dessus, insistons sur le fait que les valeurs propres de  $YX$  sont réelles strictement positives. Détaillons maintenant les calculs pour chaque ordre  $n$ .

- Soit  $n = 1$ .  $X$  et  $Y$  sont scalaires. D'après le Théorème 2.6, l'expression de  $\gamma_{\text{opt}}$  est alors trivialement obtenue par :

$$\gamma_{\text{opt}} = \sqrt{1 + YX}.$$

Reprenant les Tables 3.1-3.5 et 4.1-4.3 et en notant  $\tau := a_0/c_0$ , nous en déduisons alors :

$$\gamma_{\text{opt}}(c_0, \tau) = \sqrt{1 + \lambda_m(c_0, \tau)} = \sqrt{1 + \left(\text{sgn}(c_0)\sqrt{1 + \tau^2} - \tau\right)^2}. \quad (4.3)$$

- Soit  $n = 2$ . La Table 4.4 donne les coefficients du polynôme  $h$  de degré 2 en  $\lambda$ . Ceux-ci s'expriment en fonction de  $\beta_0$  et  $\beta_1$  définis à cette même Table 4.4. En utilisant l'Annexe D, nous montrons que les racines de celui-ci s'écrivent alors :

$$\begin{cases} \lambda_1 := \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\beta_0}}{2\Delta_o^2} > 0, \\ \lambda_2 := \frac{-\beta_1 - \sqrt{\beta_1^2 - 4\beta_0}}{2\Delta_o^2} > 0. \end{cases}$$

Comme rappelé précédemment, ces deux racines sont réelles strictement positives, ce qui implique en particulier que :

$$\beta_1^2 - 4\beta_0 > 0.$$

Trivialement, la racine réelle maximale est alors :

$$\lambda_m = \lambda_1 = \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\beta_0}}{2\Delta_o^2}.$$

Nous en déduisons l'expression sous forme close de  $\gamma_{\text{opt}}$  :

$$\gamma_{\text{opt}} = \sqrt{1 + \lambda_m}. \quad (4.4)$$

• Soit  $n = 3$ . La Table 4.4 donne les coefficients  $\nu_i, i = 0, 1, 2$ , du polynôme  $h$  de degré 3 en  $\lambda$ . Nous introduisons alors les notations suivantes :

$$\begin{cases} j := -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \\ \mu_1 := -\frac{1}{3} \nu_2^2 + \nu_1, \\ \mu_0 := \frac{2}{27} \nu_2^3 + \nu_0 - \frac{1}{3} \nu_1 \nu_2, \\ \Delta_3 := 27 (4 \mu_1^3 + 27 \mu_0^2), \\ \alpha_3 := \left( \frac{-27 \mu_0 + \sqrt{\Delta_3}}{2} \right)^{\frac{1}{3}}. \end{cases} \quad (4.5)$$

En utilisant la Proposition D.3 de l'Annexe D, nous montrons que les racines de  $h$  s'écrivent alors sous la forme de :

$$\begin{aligned} \text{Si } \mu_1 = 0, \quad & \begin{cases} \lambda_1 = (-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ \lambda_2 = j^2 (-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ \lambda_3 = j (-\mu_0)^{1/3} - \frac{\nu_2}{3}. \end{cases} \\ \text{Si } \mu_1 \neq 0, \quad & \begin{cases} \lambda_1 = \frac{1}{3} \left( \alpha_3 - \frac{3 \mu_1}{\alpha_3} \right) - \frac{\nu_2}{3}, \\ \lambda_2 = \frac{1}{3} \left( \alpha_3 - \frac{3 \mu_1}{\alpha_3} \right) j^2 - \frac{\nu_2}{3}, \\ \lambda_3 = \frac{1}{3} \left( \alpha_3 - \frac{3 \mu_1}{\alpha_3} \right) j - \frac{\nu_2}{3}. \end{cases} \end{aligned}$$

Comme rappelé précédemment, les racines de  $h$  sont toutes réelles strictement positives. Dans ce cas, le Corollaire D.1 de l'Annexe D montre que la racine réelle maximale a alors la forme suivante quelles que soient les valeurs des paramètres  $a$  et  $c$  du système :

$$\lambda_m = \lambda_1 = \frac{1}{3} \left( \alpha_3 - \frac{3 \mu_1}{\alpha_3} - \nu_2 \right).$$

Nous en déduisons alors l'expression sous forme close de  $\gamma_{\text{opt}} := \sqrt{1 + \lambda_m}$ .

• Soit  $n = 4$ . La Table 4.4 donne les coefficients du polynôme  $h$  de degré 4 en  $\lambda$ . Introduisons les notations suivantes :

$$\begin{cases} \mu_2 := -\frac{3}{8} \nu_3^2 + \nu_2, \\ \mu_1 := \frac{1}{8} \nu_3^3 - \frac{1}{2} \nu_2 \nu_3 + \nu_1, \\ \mu_0 := -\frac{3}{256} \nu_3^4 + \frac{1}{16} \nu_2 \nu_3^2 - \frac{1}{4} \nu_1 \nu_3 + \nu_0, \\ p_4 := -\left( \frac{1}{12} \mu_2^2 + \mu_0 \right), \\ q_4 := \frac{\mu_2}{3} \left( \mu_0 - \frac{\mu_2^2}{36} \right) - \frac{\mu_1^2}{8}, \end{cases} \quad \begin{cases} \delta_4 := 27 (4 p_4^3 + 27 q_4^2), \\ \alpha_4 := \left( \frac{1}{2} \left( \sqrt{\delta_4} - 27 q_4 \right) \right)^{\frac{1}{3}}, \\ u_4 := \begin{cases} \frac{1}{3} \left( \alpha_4 - \frac{3 p_4}{\alpha_4} - \mu_2 \right), & \text{si } p_4 \neq 0, \\ (-q_4)^{1/3} - \frac{\mu_2}{3}, & \text{si } p_4 = 0, \end{cases} \\ \varepsilon_1 := \pm 1, \quad \varepsilon_2 := \pm 1, \quad \varepsilon := (\varepsilon_1, \varepsilon_2) \\ \Delta_4 := -\left( u_4 + \mu_2 + \varepsilon_1 \frac{\mu_1}{\sqrt{2 u_4}} \right). \end{cases}$$

En utilisant la Proposition D.6 de l'Annexe D, nous montrons que les racines de  $h$  s'écrivent sous la forme de :

$$\lambda(\epsilon_1, \epsilon_2) := \begin{cases} \frac{\sqrt{2}}{2} \left( \epsilon_1 \sqrt{u_4} + \epsilon_2 \sqrt{\Delta_4} \right) - \frac{\nu_3}{4}, & \text{si } \mu_1 \neq 0, \\ \epsilon_1 \frac{\sqrt{2}}{2} \sqrt{\epsilon_2 \sqrt{\mu_2^2 - 4\mu_0} - \mu_2}, & \text{si } \mu_1 = 0. \end{cases}$$

D'après la Proposition C.2 de l'Annexe C, les racines de  $h$  sont toutes réelles positives. De plus, le Corollaire D.2 montre que la racine réelle maximale a alors la forme suivante quelles que soient les valeurs des paramètres :

$$\lambda_m = \lambda(1, 1) = \begin{cases} \frac{\sqrt{2}}{2} \left( \sqrt{u_4} + \sqrt{\Delta_4} \right) - \frac{\nu_3}{4}, & \text{si } \mu_1 \neq 0, \\ \frac{\sqrt{2}}{2} \sqrt{\sqrt{\mu_2^2 - 4\mu_0} - \mu_2}, & \text{si } \mu_1 = 0. \end{cases}$$

Nous en déduisons alors l'expression sous forme close de  $\gamma_{\text{opt}} := \sqrt{1 + \lambda_m}$ .

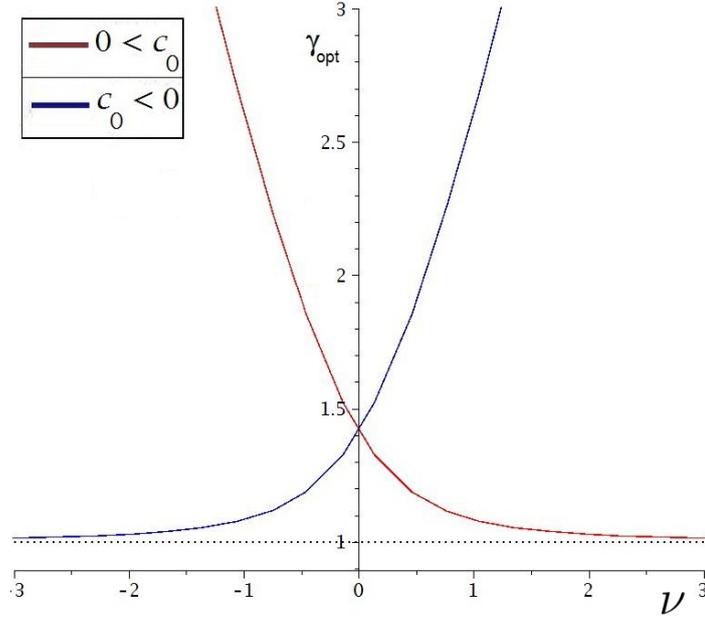
□

**Remarque 4.3.** Pour  $n = 1$ , la forme de  $\gamma_{\text{opt}}$  est très simple et son unique paramètre  $\nu = \frac{a_0}{c_0}$  permet de représenter  $\gamma_{\text{opt}}$  en fonction du signe de  $c_0$  comme le montre la Figure 4.1. Notons que si  $\nu = 0$ , c'est-à-dire si  $G = c_0/s$ , alors  $\gamma_{\text{opt}} = \sqrt{2}$  pour tout  $c_0$ . Nous retrouvons alors un résultat de [KHO07]. De plus, nous avons :

$$\lim_{\nu \rightarrow +\infty} \gamma_{\text{opt}}(c_0, \nu) = \begin{cases} 1, & c_0 > 0, \\ +\infty, & c_0 < 0, \end{cases} \quad \lim_{\nu \rightarrow -\infty} \gamma_{\text{opt}}(c_0, \nu) = \begin{cases} +\infty, & c_0 > 0, \\ 1, & c_0 < 0. \end{cases}$$

Dans les deux cas,  $\gamma_{\text{opt}}$  diverge pour  $a_0 < 0$  et  $|\nu| \rightarrow \infty$ , c'est-à-dire quand le pôle de  $G$  est instable. A contrario,  $\gamma_{\text{opt}}$  converge vers 1 pour  $a_0 > 0$  et  $|\nu| \rightarrow \infty$ , c'est-à-dire quand le pôle de  $G$  est stable.

**Remarque 4.4.** Ces formules explicites de  $\gamma_{\text{opt}}$  nous permettent alors d'en déduire la valeur du critère quelles que soient les valeurs de paramètres. Nous pouvons alors exploiter ce résultat en effectuant des tracés des variations de  $\gamma_{\text{opt}}$  par rapport aux paramètres comme montré, par exemple, dans la Figure 4.1 ou dans les exemples de la section suivante. Cela nous permet d'étudier les variations de  $\gamma_{\text{opt}}$  vis-à-vis des variations des paramètres et donc de faire une analyse des *sensibilités* du critère vis-à-vis des paramètres du système.


 FIGURE 4.1 –  $\gamma_{\text{opt}}(\nu)$  associé au système  $G = \frac{c_0}{s+a_0}$ , où  $\nu = \frac{a_0}{c_0}$ 

$n$	$\Delta_o$
1	$c_0$
2	$a_0 c_1^2 - a_1 c_0 c_1 + c_0^2$
3	$a_0^2 c_2^3 - a_0 a_1 c_1 c_2^2 - 2 a_0 a_2 c_0 c_2^2 + a_0 a_2 c_1^2 c_2 + a_1^2 c_0 c_2^2 - a_1 a_2 c_0 c_1 c_2$ $+ a_2^2 c_0^2 c_2 + 3 a_0 c_0 c_1 c_2 - a_0 c_1^3 - 2 a_1 c_0^2 c_2 + a_1 c_0 c_1^2 - a_2 c_0^2 c_1 + c_0^3$
4	$a_0^3 c_3^4 - a_0^2 a_1 c_2 c_3^3 - 2 a_0^2 a_2 c_1 c_3^3 + a_0^2 a_2 c_2^2 c_3^2 - 3 a_0^2 a_3 c_0 c_3^3 + 3 a_0^2 a_3 c_1 c_2 c_3^2 - a_0^2 a_3 c_2^2 c_3$ $+ a_0 a_1^2 c_1 c_3^3 + 3 a_0 a_1 a_2 c_0 c_3^3 - a_0 a_1 a_2 c_1 c_2 c_3^2 - a_0 a_1 a_3 c_0 c_2 c_3^2 - 2 a_0 a_1 a_3 c_1^2 c_3^2$ $+ a_0 a_1 a_3 c_1 c_2^2 c_3 - 2 a_0 a_2^2 c_0 c_2 c_3^2 + a_0 a_2^2 c_1^2 c_3^2 + a_0 a_2 a_3 c_0 c_1 c_3^2 + 2 a_0 a_2 a_3 c_0 c_2^2 c_3$ $- a_0 a_2 a_3 c_1^2 c_2 c_3 + 3 a_0 a_3^2 c_0^2 c_3^2 - 3 a_0 a_3^2 c_0 c_1 c_2 c_3 + a_0 a_3^2 c_1^3 c_3 - a_1^3 c_0 c_3^3$ $+ a_1^2 a_2 c_0 c_2 c_3^2 + 2 a_1^2 a_3 c_0 c_1 c_3^2 - a_1^2 a_3 c_0 c_2^2 c_3 - a_1 a_2^2 c_0 c_1 c_3^2 - 3 a_1 a_2 a_3 c_0^2 c_3^2$ $+ a_1 a_2 a_3 c_0 c_1 c_2 c_3 + 2 a_1 a_3^2 c_0^2 c_2 c_3 - a_1 a_3^2 c_0 c_1^2 c_3 + a_2^3 c_0^2 c_3^2 - a_2^2 a_3 c_0^2 c_2 c_3$ $+ a_2 a_3^2 c_0^2 c_1 c_3 - a_3^3 c_0^3 c_3 + 4 a_0^2 c_0 c_2 c_3^2 + 2 a_0^2 c_1^2 c_3^2 - 4 a_0^2 c_1 c_2^2 c_3 + a_0 c_1^4$ $+ a_0^2 c_2^4 - 5 a_0 a_1 c_0 c_1 c_3^2 + a_0 a_1 c_0 c_2^2 c_3 + 3 a_0 a_1 c_1^2 c_2 c_3 - a_0 a_1 c_1 c_2^3 - 3 a_0 a_2 c_0^2 c_3^2$ $+ 4 a_0 a_2 c_0 c_1 c_2 c_3 - 2 a_0 a_2 c_0 c_2^3 - 2 a_0 a_2 c_1^3 c_3 + a_0 a_2 c_1^2 c_2^2 - 5 a_0 a_3 c_0^2 c_2 c_3 + a_0 a_3 c_0 c_1^2 c_3$ $+ 3 a_0 a_3 c_0 c_1 c_2^2 - a_0 a_3 c_1^3 c_2 + 3 a_1^2 c_0^2 c_3^2 - 3 a_1^2 c_0 c_1 c_2 c_3 + a_1^2 c_0 c_2^3 + a_1 a_2 c_0^2 c_2 c_3$ $+ 2 a_1 a_2 c_0 c_1^2 c_3 - a_1 a_2 c_0 c_1 c_2^2 - a_1 a_3 c_0^2 c_1 c_3 - 2 a_1 a_3 c_0^2 c_2^2 + a_1 a_3 c_0 c_1^2 c_2 - 2 a_2^2 c_0^2 c_1 c_3$ $+ a_2^2 c_0^2 c_2^2 + 3 a_2 a_3 c_0^3 c_3 - a_2 a_3 c_0^2 c_1 c_2 + a_2^3 c_0^3 c_2 + 4 a_0 c_0^2 c_1 c_3 + 2 a_0 c_0^2 c_2^2$ $- 4 a_0 c_0 c_1^2 c_2 - 3 a_1 c_0^3 c_3 + 3 a_1 c_0^2 c_1 c_2 - a_1 c_0 c_1^3 - 2 a_2 c_0^3 c_2 + a_2 c_0^2 c_1^2 - a_3 c_0^3 c_1 + c_0^4$

 TABLE 4.1 – Conditions de non-observabilité  $\Delta_o = 0$  pour  $1 \leq n \leq 4$

$n$	Matrice $\mathcal{Q}$
1	$\mathcal{Q} := \Delta_o^{-1} = c_0^{-1}$
2	$\mathcal{Q} := \frac{1}{\Delta_o} \begin{pmatrix} q_1 & q_2 \\ q_2 & q_3 \end{pmatrix}$ $q_1 := -c_1, \quad q_2 := c_0, \quad q_3 := c_1 a_0 - a_1 c_0$
3	$\mathcal{Q} := \frac{1}{\Delta_o} \begin{pmatrix} q_1 & q_2 & q_3 \\ q_2 & q_3 & q_4 \\ q_3 & q_4 & q_5 \end{pmatrix}$ $q_1 := a_1 c_2^2 - (a_2 c_1 + c_0) c_2 + c_1^2$ $q_2 := -a_0 c_2^2 + a_2 c_0 c_2 - c_0 c_1$ $q_3 := a_0 c_1 c_2 - a_1 c_0 c_2 + c_0^2$ $q_4 := a_0 c_0 c_2 - a_0 c_1^2 + a_1 c_0 c_1 - a_2 c_0^2$ $q_5 := (a_0 c_2 - a_2 c_0)^2 - a_0 a_1 c_1 c_2 + a_0 a_2 c_1^2 + a_1^2 c_0 c_2 - a_1 a_2 c_0 c_1 + a_0 c_0 c_1 - a_1 c_0^2$
4	$\mathcal{Q} := \frac{1}{\Delta_o} \begin{pmatrix} q_1 & q_2 & q_3 & q_4 \\ q_2 & q_3 & q_4 & q_5 \\ q_3 & q_4 & q_5 & q_6 \\ q_4 & q_5 & q_6 & q_7 \end{pmatrix}$ $q_1 := a_0 a_2 c_3^3 - a_0 a_3 c_2 c_3^2 - a_1^2 c_3^3 + a_1 a_2 c_2 c_3^2 + 2 a_1 a_3 c_1 c_3^2 - a_1 a_3 c_2^2 c_3 - a_2^2 c_1 c_3^2$ $+ a_1 c_2^3 - a_2 a_3 c_0 c_3^2 + a_2 a_3 c_1 c_2 c_3 + a_3^2 c_0 c_2 c_3 - a_3^2 c_1^2 c_3 - a_0 c_1 c_3^2 + a_0 c_2^2 c_3$ $+ 2 a_1 c_0 c_3^2 - 3 a_1 c_1 c_2 c_3 + 2 a_2 c_1^2 c_3 - a_2 c_1 c_2^2 - a_3 c_0 c_1 c_3 - a_3 c_0 c_2^2 + a_3 c_1^2 c_2$ $- c_0^2 c_3 + 2 c_0 c_1 c_2 - c_1^3$ $q_2 := a_0 a_1 c_3^3 - a_0 a_2 c_2 c_3^2 - a_0 a_3 c_1 c_3^2 + a_0 a_3 c_2^2 c_3 - a_1 a_3 c_0 c_3^2 + a_2^2 c_0 c_3^2 - a_2 a_3 c_0 c_2 c_3$ $+ a_3^2 c_0 c_1 c_3 - a_0 c_0 c_3^2 + 2 a_0 c_1 c_2 c_3 - a_0 c_2^3 + a_1 c_0 c_2 c_3 - 2 a_2 c_0 c_1 c_3 + a_2 c_0 c_2^2 + a_3 c_0^2 c_3$ $- a_3 c_0 c_1 c_2 - c_0^2 c_2 + c_0 c_1^2$ $q_3 := -a_0^2 c_3^3 + a_0 a_2 c_1 c_3^2 + 2 a_0 a_3 c_0 c_3^2 - a_0 a_3 c_1 c_2 c_3 - a_1 a_2 c_0 c_3^2 + a_1 a_3 c_0 c_2 c_3 - a_3^2 c_0^2 c_3$ $- 2 a_0 c_0 c_2 c_3 - a_0 c_1^2 c_3 + a_0 c_1 c_2^2 + a_1 c_0 c_1 c_3 - a_1 c_0 c_2^2 + a_2 c_0^2 c_3 + a_3 c_0^2 c_2 - c_0^2 c_1$ $q_4 := a_0^2 c_2 c_3^2 - a_0 a_1 c_1 c_3^2 - a_0 a_2 c_0 c_3^2 - a_0 a_3 c_0 c_2 c_3 + a_0 a_3 c_1^2 c_3 + a_1^2 c_0 c_3^2 - a_1 a_3 c_0 c_1 c_3$ $+ a_2 a_3 c_0^2 c_3 + 2 a_0 c_0 c_1 c_3 + a_0 c_0 c_2^2 - a_0 c_1^2 c_2 - 2 a_1 c_0^2 c_3 + a_1 c_0 c_1 c_2 - a_2 c_0^2 c_2 + c_0^3$ $q_5 := a_0^2 c_1 c_3^2 - a_0^2 c_2^2 c_3 - a_0 a_1 c_0 c_3^2 + a_0 a_1 c_1 c_2 c_3 + 2 a_0 a_2 c_0 c_2 c_3 - a_0 a_2 c_1^2 c_3 - a_0 a_3 c_0 c_1 c_3$ $- a_1^2 c_0 c_2 c_3 + a_1 a_2 c_0 c_1 c_3 + a_1 a_3 c_0^2 c_3 - a_2^2 c_0^2 c_3 + a_0 c_0^2 c_3 - 2 a_0 c_0 c_1 c_2 + a_0 c_1^3$ $+ a_1 c_0^2 c_2 - a_1 c_0 c_1^2 + a_2 c_0^2 c_1 - a_3 c_0^3$ $q_6 := a_0^2 c_0 c_3^2 - 2 a_0^2 c_1 c_2 c_3 + a_0^2 c_2^3 + a_0 a_1 c_0 c_2 c_3 + a_0 a_1 c_1^2 c_3 - a_0 a_1 c_1 c_2^2 - 2 a_0 a_2 c_0 c_2^2$ $+ a_0 a_2 c_1^2 c_2 - 2 a_0 a_3 c_0^2 c_3 + 3 a_0 a_3 c_0 c_1 c_2 - a_0 a_3 c_1^3 - a_1^2 c_0 c_1 c_3 + a_1^2 c_0 c_2^2$ $+ a_1 a_2 c_0^2 c_3 - a_1 a_2 c_0 c_1 c_2 - 2 a_1 a_3 c_0^2 c_2 + a_1 a_3 c_0 c_1^2 + a_2^2 c_0^2 c_2 - a_2 a_3 c_0^2 c_1$ $+ a_3^2 c_0^3 + a_0 c_0^2 c_2 - a_0 c_0 c_1^2 + a_1 c_0^2 c_1 - a_2 c_0^3$ $q_7 := a_0^3 c_3^3 - a_0^2 a_1 c_2 c_3^2 - 2 a_0^2 a_2 c_1 c_3^2 + a_0^2 a_2 c_2^2 c_3 - 3 a_0^2 a_3 c_0 c_3^2 + 3 a_0^2 a_3 c_1 c_2 c_3 - a_0^2 a_3 c_2^3$ $+ a_0 a_1^2 c_1 c_3^2 + 3 a_0 a_1 a_2 c_0 c_3^2 - a_0 a_1 a_2 c_1 c_2 c_3 - a_0 a_1 a_3 c_0 c_2 c_3 - 2 a_0 a_1 a_3 c_1^2 c_3$ $+ a_0 a_1 a_3 c_1 c_2^2 - 2 a_0 a_2^2 c_0 c_2 c_3 + a_0 a_2^2 c_1^2 c_3 + a_0 a_2 a_3 c_0 c_1 c_3 + 2 a_0 a_2 a_3 c_0 c_2^2$ $+ 3 a_0 a_3^2 c_0^2 c_3 - 3 a_0 a_3^2 c_0 c_1 c_2 + a_0 a_3^2 c_1^3 - a_1^3 c_0 c_3^2 + a_1^2 a_2 c_0 c_2 c_3 + 2 a_1^2 a_3 c_0 c_1 c_3$ $- a_1^2 a_3 c_0 c_2^2 + a_2^3 c_0^2 c_3 - a_1 a_2^2 c_0 c_1 c_3 - 3 a_1 a_2 a_3 c_0^2 c_3 + a_1 a_2 a_3 c_0 c_1 c_2 + 2 a_1 a_3^2 c_0^2 c_2$ $- a_1 a_3^2 c_0 c_1^2 - a_2^2 a_3 c_0^2 c_2 + a_2 a_3^2 c_0^2 c_1 - a_3^3 c_0^3 + 2 a_0^2 c_0 c_2 c_3 + a_0^2 c_1^2 c_3 - a_0^2 c_1 c_2^2$ $- 3 a_0 a_1 c_0 c_1 c_3 + a_0 a_1 c_1^2 c_2 - 2 a_0 a_2 c_0^2 c_3 + 2 a_0 a_2 c_0 c_1 c_2 - a_0 a_2 c_1^3 - 2 a_0 a_3 c_0^2 c_2$ $+ 2 a_1^2 c_0^2 c_3 - a_1^2 c_0 c_1 c_2 + a_1 a_2 c_0 c_1^2 - a_1 a_3 c_0^2 c_1 - a_2^2 c_0^2 c_1 + 2 a_2 a_3 c_0^3 + a_0 c_0^2 c_1$ $+ a_0 a_3 c_0 c_1^2 - a_0 a_2 a_3 c_1^2 c_2 - a_1 c_0^3$

TABLE 4.2 – Matrices  $\mathcal{Q}$  pour  $1 \leq n \leq 4$

$n$	Matrice $Y$
1	$Y := \left( \sqrt{a_0^2 + c_0^2} - a_0 \right) c_0^{-2}$
2	$Y := \frac{1}{\Delta_o^2} \begin{pmatrix} z_{1,1} & z_{1,2} \\ z_{1,2} & z_{2,2} \end{pmatrix}$ $z_{1,1} := q_1^2 x_{1,1} + 2 q_1 q_2 x_{1,2} + q_2^2 x_{2,2}$ $z_{1,2} := (q_2 x_{1,1} + q_3 x_{1,2}) q_1 + (q_2 x_{1,2} + q_3 x_{2,2}) q_2$ $z_{2,2} := q_2^2 x_{1,1} + 2 q_2 q_3 x_{1,2} + q_3^2 x_{2,2}$
3	$Y := \frac{1}{\Delta_o^2} \begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} \\ z_{1,2} & z_{2,2} & z_{2,3} \\ z_{1,3} & z_{2,3} & z_{3,3} \end{pmatrix}$ $z_{1,1} := q_1^2 x_{1,1} + q_2^2 x_{2,2} + q_3^2 x_{3,3} + 2 (x_{1,2} q_1 q_2 + x_{1,3} q_1 q_3 + x_{2,3} q_2 q_3)$ $z_{1,2} := q_2^2 x_{1,2} + q_3^2 x_{2,3} + (q_1 x_{1,2} + q_4 x_{3,3}) q_3 + ((x_{1,3} + x_{2,2}) q_3 + q_1 x_{1,1} + q_4 x_{2,3}) q_2 + q_1 q_4 x_{1,3}$ $z_{1,3} := q_3^2 x_{1,3} + q_1 (q_4 x_{1,2} + q_5 x_{1,3}) + q_2 (q_4 x_{2,2} + q_5 x_{2,3}) + (q_1 x_{1,1} + q_2 x_{1,2} + q_4 x_{2,3} + q_5 x_{3,3}) q_3$ $z_{2,2} := q_2^2 x_{1,1} + q_3^2 x_{2,2} + q_4^2 x_{3,3} + 2 (q_2 x_{1,3} + q_3 x_{2,3}) q_4 + 2 q_2 q_3 x_{1,2}$ $z_{2,3} := q_3^2 x_{1,2} + q_4^2 x_{2,3} + (q_2 x_{1,2} + q_5 x_{3,3}) q_4 + ((x_{1,3} + x_{2,2}) q_4 + q_2 x_{1,1} + q_5 x_{2,3}) q_3 + q_2 q_5 x_{1,3}$ $z_{3,3} := q_3^2 x_{1,1} + q_4^2 x_{2,2} + q_5^2 x_{3,3} + 2 q_3 q_4 x_{1,2} + 2 (q_3 x_{1,3} + q_4 x_{2,3}) q_5$
4	$Y := \frac{1}{\Delta_o^2} \begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} & z_{1,4} \\ z_{1,2} & z_{2,2} & z_{2,3} & z_{2,4} \\ z_{1,3} & z_{2,3} & z_{3,3} & z_{3,4} \\ z_{1,4} & z_{2,4} & z_{3,4} & z_{4,4} \end{pmatrix}$ $z_{1,1} := q_1^2 x_{1,1} + (2 q_2 x_{1,2} + 2 q_3 x_{1,3} + 2 q_4 x_{1,4}) q_1 + q_2^2 x_{2,2} + (2 q_3 x_{2,3} + 2 q_4 x_{2,4}) q_2 + x_{3,3} q_3^2 + 2 x_{3,4} q_3 q_4 + x_{4,4} q_4^2$ $z_{1,2} := q_2^2 x_{1,2} + ((x_{2,2} + x_{1,3}) q_3 + (x_{1,4} + x_{2,3}) q_4 + q_1 x_{1,1} + q_5 x_{2,4}) q_2 + x_{2,3} q_3^2 + ((x_{2,4} + x_{3,3}) q_4 + q_1 x_{1,2} + q_5 x_{3,4}) q_3 + x_{3,4} q_4^2 + (q_1 x_{1,3} + q_5 x_{4,4}) q_4 + q_1 q_5 x_{1,4}$ $z_{1,3} := q_3^2 x_{1,3} + ((x_{1,4} + x_{2,3}) q_4 + q_1 x_{1,1} + q_5 x_{3,3} + q_6 x_{3,4} + q_2 x_{1,2}) q_3 + x_{2,4} q_4^2 + (q_1 x_{1,2} + q_2 x_{2,2} + q_5 x_{3,4} + q_6 x_{4,4}) q_4 + (q_5 x_{1,3} + q_6 x_{1,4}) q_1 + q_2 (q_5 x_{2,3} + q_6 x_{2,4})$ $z_{1,4} := x_{1,4} q_4^2 + (q_1 x_{1,1} + q_2 x_{1,2} + q_3 x_{1,3} + q_5 x_{2,4} + q_6 x_{3,4} + q_7 x_{4,4}) q_4 + (q_5 x_{1,2} + q_6 x_{1,3} + q_7 x_{1,4}) q_1 + (q_5 x_{2,2} + q_6 x_{2,3} + q_7 x_{2,4}) q_2 + q_3 (q_5 x_{2,3} + q_6 x_{3,3} + q_7 x_{3,4})$ $z_{2,2} := q_2^2 x_{1,1} + (2 q_3 x_{1,2} + 2 q_4 x_{1,3} + 2 q_5 x_{1,4}) q_2 + q_3^2 x_{2,2} + (2 q_4 x_{2,3} + 2 q_5 x_{2,4}) q_3 + x_{3,3} q_4^2 + 2 x_{3,4} q_4 q_5 + x_{4,4} q_5^2$ $z_{2,3} := q_3^2 x_{1,2} + ((x_{2,2} + x_{1,3}) q_4 + (x_{1,4} + x_{2,3}) q_5 + q_2 x_{1,1} + q_6 x_{2,4}) q_3 + x_{2,3} q_4^2 + ((x_{2,4} + x_{3,3}) q_5 + q_2 x_{1,2} + q_6 x_{3,4}) q_4 + x_{3,4} q_5^2 + (q_2 x_{1,3} + q_6 x_{4,4}) q_5 + q_2 q_6 x_{1,4}$ $z_{2,4} := q_4^2 x_{1,3} + ((x_{1,4} + x_{2,3}) q_5 + q_2 x_{1,1} + q_6 x_{3,3} + q_7 x_{3,4} + q_3 x_{1,2}) q_4 + x_{2,4} q_5^2 + (q_2 x_{1,2} + q_3 x_{2,2} + q_6 x_{3,4} + q_7 x_{4,4}) q_5 + (q_6 x_{1,3} + q_7 x_{1,4}) q_2 + q_3 (q_6 x_{2,3} + q_7 x_{2,4})$ $z_{3,3} := q_3^2 x_{1,1} + (2 q_4 x_{1,2} + 2 q_5 x_{1,3} + 2 q_6 x_{1,4}) q_3 + q_4^2 x_{2,2} + (2 q_5 x_{2,3} + 2 q_6 x_{2,4}) q_4 + x_{3,3} q_5^2 + 2 x_{3,4} q_5 q_6 + x_{4,4} q_6^2$ $z_{3,4} := q_4^2 x_{1,2} + ((x_{2,2} + x_{1,3}) q_5 + (x_{1,4} + x_{2,3}) q_6 + q_3 x_{1,1} + q_7 x_{2,4}) q_4 + x_{2,3} q_5^2 + ((x_{2,4} + x_{3,3}) q_6 + q_3 x_{1,2} + q_7 x_{3,4}) q_5 + x_{3,4} q_6^2 + (q_3 x_{1,3} + q_7 x_{4,4}) q_6 + q_3 q_7 x_{1,4}$ $z_{4,4} := q_4^2 x_{1,1} + (2 q_5 x_{1,2} + 2 q_6 x_{1,3} + 2 q_7 x_{1,4}) q_4 + q_5^2 x_{2,2} + (2 q_6 x_{2,3} + 2 q_7 x_{2,4}) q_5 + x_{3,3} q_6^2 + 2 x_{3,4} q_6 q_7 + x_{4,4} q_7^2$

TABLE 4.3 – Matrices  $Y$  pour  $1 \leq n \leq 4$

$n$	Polynôme $h$
1	$h(\lambda) := \lambda - YX = \lambda - \left( \sqrt{a_0^2 + c_0^2} - a_0 \right)^2 c_0^{-2}$
2	$h(\lambda) := \lambda^2 + \nu_1 \lambda + \nu_0$ $\nu_1 := \beta_1(a, c, b) \Delta_o^{-2}$ $\nu_0 := \beta_0(a, c, b) \Delta_o^{-4} = (\det(X) \Delta_o^{-1})^2$ $\beta_1 := -(z_{1,1} x_{1,1} + z_{2,2} x_{2,2} + 2 z_{1,2} x_{1,2})$ $\beta_0 := (\Delta_o \det(X))^2$
3	$h(\lambda) := \lambda^3 + \nu_2 \lambda^2 + \nu_1 \lambda + \nu_0$ $\nu_k := \beta_k \Delta_o^{2(k-n)}, n = 3, k = 0, 1, 2$ $\beta_2 := -(x_{1,1} z_{1,1} + x_{2,2} z_{2,2} + x_{3,3} z_{3,3} + 2(x_{1,2} z_{1,2} + x_{1,3} z_{1,3} + x_{2,3} z_{2,3}))$ $\beta_1 := (z_{1,2}^2 - z_{1,1} z_{2,2}) x_{1,2}^2 + (z_{1,3}^2 - z_{1,1} z_{3,3}) x_{1,3}^2 + (z_{2,3}^2 - z_{2,2} z_{3,3}) x_{2,3}^2 - x_{1,1} x_{2,2} z_{1,2}^2$ $- x_{1,1} x_{3,3} z_{1,3}^2 - x_{2,2} x_{3,3} z_{2,3}^2 + z_{1,1} z_{3,3} x_{1,1} x_{3,3} + z_{1,1} z_{2,2} x_{1,1} x_{2,2} + z_{2,2} z_{3,3} x_{2,2} x_{3,3}$ $+ 2(x_{1,3}(z_{1,2} z_{1,3} - z_{1,1} z_{2,3}) + x_{2,3}(z_{1,2} z_{2,3} - z_{1,3} z_{2,2}) + x_{3,3}(z_{1,2} z_{3,3} - z_{1,3} z_{2,3})) x_{1,2}$ $+ 2[x_{2,3}(z_{1,3} z_{2,3} - z_{1,2} z_{3,3}) + x_{2,2}(z_{1,3} z_{2,2} - z_{1,2} z_{2,3})] x_{1,3}$ $+ 2x_{1,1} x_{2,3}(z_{1,1} z_{2,3} - z_{1,2} z_{1,3})$ $\beta_0 := -\Delta_o^4 \det(X)^2$
4	$h(\lambda) := \lambda^4 + \nu_3 \lambda^3 + \nu_2 \lambda^2 + \nu_1 \lambda + \nu_0$ Les coefficients $\nu_k$ sont trop longs pour être affichés ici

 TABLE 4.4 – Polynôme caractéristique  $h$  de la matrice  $YX$  pour  $1 \leq n \leq 4$

$n$	$\gamma_{\text{opt}}$
1	$\nu := \frac{a_0}{c_0}$ $\gamma_{\text{opt}} = \sqrt{1 + (\text{sgn}(c_0)\sqrt{1 + \nu^2} - \nu)^2}$
2	$\lambda_m := \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\beta_0}}{2\Delta_0^2}$ $\gamma_{\text{opt}} = \sqrt{1 + \lambda_m}$
3	$j := -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ $\mu_1 := -\frac{1}{3}\nu_2^2 + \nu_1$ $\mu_0 := \frac{2}{27}\nu_2^3 + \nu_0 - \frac{1}{3}\nu_1\nu_2$ $\Delta_3 := 27(4\mu_1^3 + 27\mu_0^2)$ $\alpha_3 := \left(\frac{-27\mu_0 + \sqrt{\Delta_3}}{2}\right)^{\frac{1}{3}}$ $\lambda_m := \frac{1}{3}\left(\alpha_3 - \frac{3\mu_1}{\alpha_3} - \nu_2\right)$ $\gamma_{\text{opt}} = \sqrt{1 + \lambda_m}$
4	$\mu_2 := -\frac{3}{8}\nu_3^2 + \nu_2$ $\mu_1 := \frac{1}{8}\nu_3^3 - \frac{1}{2}\nu_2\nu_3 + \nu_1$ $\mu_0 := -\frac{3}{256}\nu_3^4 + \frac{1}{16}\nu_2\nu_3^2 - \frac{1}{4}\nu_1\nu_3 + \nu_0$ $p_4 := -\left(\frac{1}{12}\mu_2^2 + \mu_0\right)$ $q_4 := \frac{\mu_2}{3}\left(\mu_0 - \frac{\mu_2^2}{36}\right) - \frac{\mu_1^2}{8}$ $\delta_4 := 27(4p_4^3 + 27q_4^2)$ $\alpha_4 := \left(\frac{1}{2}\left(\sqrt{\delta_4} - 27q_4\right)\right)^{\frac{1}{3}}$ $u_4 := \begin{cases} \frac{1}{3}\left(\alpha_4 - \frac{3p_4}{\alpha_4} - \mu_2\right), & \text{si } p_4 \neq 0 \\ (-q_4)^{1/3} - \frac{\mu_2}{3}, & \text{si } p_4 = 0 \end{cases}$ $\varepsilon_1 := \pm 1, \quad \varepsilon_2 := \pm 1, \quad \varepsilon := (\varepsilon_1, \varepsilon_2)$ $\Delta_4 := -\left(u_4 + \mu_2 + \varepsilon_1 \frac{\mu_1}{\sqrt{2}u_4}\right)$ $\lambda_m := \begin{cases} \frac{\sqrt{2}}{2}\left(\sqrt{u_4} + \sqrt{\Delta_4}\right) - \frac{\nu_3}{4}, & \text{si } \mu_1 \neq 0 \\ \frac{\sqrt{2}}{2}\sqrt{\sqrt{\mu_2^2 - 4\mu_0} - \mu_2}, & \text{si } \mu_1 = 0 \end{cases}$ $\gamma_{\text{opt}} = \sqrt{1 + \lambda_m}$

 TABLE 4.5 – Formes closes de  $\gamma_{\text{opt}}$  pour  $1 \leq n \leq 4$

### 4.1.2 Correcteurs $H_\infty$ paramétriques

Disposant explicitement de  $\gamma_{\text{opt}}$  en fonction des paramètres pour des systèmes d'ordre au plus 4, l'application du Corollaire 2.1 pour  $\gamma > \gamma_{\text{opt}}$  permet alors de déduire un correcteur  $H_\infty$  explicite noté  $K_\gamma$ .

En particulier, pour des systèmes à faible nombre de paramètres, nous montrons que les méthodes de calcul formel permettent d'expliciter le correcteur  $H_\infty$  optimal, c'est-à-dire obtenu pour la valeur limite  $\gamma = \gamma_{\text{opt}}$ , dépendant uniquement des paramètres du système. Nous noterons ce correcteur optimal  $K_{\gamma_{\text{opt}}}$ .

Dans un souci de clarté, nous choisissons de montrer ici quelques correcteurs  $K_\gamma$  pour des systèmes généraux d'ordre 1 et 2. Le lecteur comprendra aisément pourquoi de tels correcteurs peuvent s'avérer complexes à afficher, voir même à calculer, pour des systèmes comportant plus de paramètres.

- Pour  $n = 1$ , comme  $BB^T = 1$  et  $YC^TC = X$ , le Corollaire 2.1 donne alors :

$$\begin{cases} Z_\gamma = (\gamma_{\text{opt}}^2 - \gamma^2)^{-1}, \\ A_\gamma = -a_0 - |c_0| \sqrt{\gamma_{\text{opt}}^2 - 1} \frac{\gamma_{\text{opt}}^2 - 2\gamma^2}{\gamma_{\text{opt}}^2 - \gamma^2}, \\ B_\gamma = -\frac{\text{sgn}(c_0) \gamma^2}{\gamma_{\text{opt}}^2 - \gamma^2} \sqrt{\gamma_{\text{opt}}^2 - 1}, \\ C_\gamma = |c_0| \sqrt{\gamma_{\text{opt}}^2 - 1}. \end{cases}$$

Nous obtenons alors un ensemble de correcteurs (sous-)optimaux, paramétrés par  $\gamma \geq \gamma_{\text{opt}}$ , de fonction de transfert suivante :

$$\gamma \geq \gamma_{\text{opt}}, \quad K_\gamma(s) = \frac{-c_0 \gamma^2 (\gamma_{\text{opt}}^2 - 1)}{(s + a_0)(\gamma_{\text{opt}}^2 - \gamma^2) + |c_0| (\gamma_{\text{opt}}^2 - 2\gamma^2) \sqrt{\gamma_{\text{opt}}^2 - 1}}. \quad (4.6)$$

Pour  $\gamma = \gamma_{\text{opt}}$ , nous remarquons une simplification dans (4.6) et le correcteur  $H_\infty$  optimal a finalement la forme suivante (simple gain) :

$$K_{\gamma_{\text{opt}}} = \text{sgn}(c_0) \sqrt{\gamma_{\text{opt}}^2 - 1}.$$

Remarquons que  $K_{\gamma_{\text{opt}}}$  est d'ordre 0, tandis que  $K_\gamma$  est d'ordre 1. De telles réductions d'ordre du correcteur se vérifient à n'importe quel  $n$ , voir [DF00] ou [ACA<sup>+</sup>99, Th. 4.2.2]. Celle-ci s'explique par la singularité de la matrice  $Z_{\gamma_{\text{opt}}}$  (voir Corollaire 2.1).

- Pour  $n = 2$ , le calcul de  $K_\gamma$  est aussi réalisable, mais commence déjà à être difficilement affichable sur papier. Celui-ci a la forme suivante

$$K(s) := \frac{N_K(s)}{D_K(s)},$$

où :

$$\begin{aligned}
 D_K(s) := & \left( (-\gamma^4 + 8\gamma^2 - 8) c_0^4 + 2c_1 \left( (\gamma^4 - 4) a_1 - 3\gamma^2 b_1 + 4b_1 \right) c_0^3 + \left( (2\gamma^2 - 4) c_1^4 \right. \right. \\
 & + \left( (-\gamma^4 + 4\gamma^2 - 8) a_1^2 + 4a_1 b_1 + (-2\gamma^4 - 4\gamma^2 + 8) a_0 + 8\gamma^2 b_0 - 12b_0 \right) c_1^2 + \left( 2\gamma^2 - 4 \right) a_1^4 \\
 & + \left( -2\gamma^2 b_1 + 4b_1 \right) a_1^3 + \left( -4\gamma^2 a_0 + 8\gamma^2 b_0 - 12b_0 \right) a_1^2 + 2b_1 \left( \gamma^2 a_0 - 3\gamma^2 b_0 + 4b_0 \right) a_1 \\
 & + 8a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) c_0^2 + 8c_1 \left( \left( (1/4\gamma^4 - 1/2\gamma^2) a_0 - 1/2b_0 \right) a_1 \right. \\
 & - 1/4b_0 b_1 (\gamma^2 - 2) \left. \right) c_1^2 + \left( -1/2b_0 - 1/2\gamma^2 a_0 \right) a_1^3 + 1/2b_1 \left( \gamma^2 a_0 + b_0 \right) a_1^2 \\
 & + a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) a_1 - a_0 (\gamma - 1) (\gamma + 1) b_1 (-b_0 + a_0) c_0 \\
 & - 8a_0 \left( 1/8 (\gamma^2 - 2)^2 a_0 c_1^4 + \left( (1/2b_0 + a_0 - 1/2\gamma^2 a_0) a_1^2 + 1/4b_1 (\gamma^2 - 2) (a_0 + b_0) a_1 \right. \right. \\
 & \left. \left. + a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) \right) c_1^2 + 1/2a_1^3 (a_1 - b_1) (a_0 + b_0) \right) s^2 \\
 & + \left( \left( (\gamma^4 - 16\gamma^2 + 16) a_1 - 2b_1\gamma^4 + 10\gamma^2 b_1 - 8b_1 \right) c_0^4 + 2c_1 \left( -8 (\gamma - 1) (\gamma + 1) c_1^2 \right. \right. \\
 & + \left( -\gamma^4 - 4\gamma^2 + 8 \right) a_1^2 + \left( 2b_1\gamma^4 + \gamma^2 b_1 - 4b_1 \right) a_1 + 16\gamma^2 a_0 - 8\gamma^2 b_0 - 16a_0 + 8b_0 \left. \right) c_0^3 \\
 & + \left( \left( (-2\gamma^2 + 4) a_1 + 4\gamma^2 b_1 - 4b_1 \right) c_1^4 + \left( (\gamma^4 - 4\gamma^2 + 8) a_1^3 - 2b_1 (\gamma^4 - 3\gamma^2 + 4) a_1^2 \right. \right. \\
 & + \left( (2\gamma^4 - 4\gamma^2) a_0 - 12\gamma^2 b_0 + 16b_0 \right) a_1 - 4 (\gamma - 1) \left( (\gamma^2 + 2) a_0 - 3b_0 \right) (\gamma + 1) b_1 \left. \right) c_1^2 \\
 & + \left( -2\gamma^2 + 4 \right) a_1^5 + \left( 2\gamma^2 b_1 - 4b_1 \right) a_1^4 + \left( 4\gamma^2 a_0 - 12\gamma^2 b_0 + 16b_0 \right) a_1^3 \\
 & - 2b_1 \left( \gamma^2 a_0 - 5\gamma^2 b_0 + 6b_0 \right) a_1^2 - 16a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) a_1 \\
 & + 8a_0 (\gamma - 1) (\gamma + 1) b_1 (-b_0 + a_0) c_0^2 + 8c_1 \left( -1/2b_0 (\gamma - 1) (\gamma + 1) c_1^4 \right. \\
 & + \left( -1/4 (\gamma^2 - 2) \left( \gamma^2 a_0 + 2b_0 \right) a_1^2 + 1/2 \left( (\gamma^4 - \gamma^2) a_0 + 1/2\gamma^2 b_0 - b_0 \right) b_1 a_1 \right. \\
 & \left. - 2a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) \right) c_1^2 + 1/2 \left( \gamma^2 a_0 + b_0 \right) a_1^4 - 1/2b_1 \left( \gamma^2 a_0 + b_0 \right) a_1^3 \\
 & - 2a_0 (\gamma - 1) (\gamma + 1) (-b_0 + a_0) a_1^2 + a_0 (\gamma - 1) (\gamma + 1) b_1 (-b_0 + a_0) a_1 \\
 & + 4 (\gamma - 1) (\gamma + 1) a_0^2 (-b_0 + a_0) c_0 - 8a_0 \left( \left( -1/8 \left( (\gamma^2 - 2)^2 a_0 - 4\gamma^2 b_0 + 4b_0 \right) a_1 \right. \right. \\
 & + 1/8 \left( 2b_1\gamma^4 - 6\gamma^2 b_1 + 4b_1 \right) a_0 \left. \right) c_1^4 + \left( 1/2 (\gamma^2 - 2) (a_0 + b_0) a_1^3 - 3/4 \left( (\gamma^2 - 4/3) a_0 \right. \right. \\
 & \left. \left. + 1/3b_0 (\gamma^2 - 2) \right) b_1 a_1^2 + a_0 (\gamma - 1) (\gamma + 1) b_1 (-b_0 + a_0) \right) c_1^2 - 1/2a_1^4 (-b_1 + a_1) (a_0 + b_0) \left. \right) s \\
 & + \left( (-2\gamma^4 + 14\gamma^2 - 12) c_1^2 + (-2\gamma^4 + 14\gamma^2 - 12) a_1^2 + \left( 2b_1\gamma^4 - 10\gamma^2 b_1 + 8b_1 \right) a_1 + \left( 3\gamma^4 - 12\gamma^2 + 8 \right) a_0 - 4b_0\gamma^4 + 12\gamma^2 b_0 - 8b_0 \right) c_0^4 \\
 & + 2c_1 \left( -8 (\gamma - 1) \left( (-1/8\gamma^2 - 1/4) a_1 + 1/4b_1 \right) (\gamma + 1) c_1^2 + \left( \gamma^4 + \gamma^2 - 2 \right) a_1^3 \right. \\
 & + \left( -b_1\gamma^4 - \gamma^2 b_1 + 2b_1 \right) a_1^2 + \left( (-\gamma^4 - 6\gamma^2 + 4) a_0 + 2b_0\gamma^4 + 2\gamma^2 b_0 - 4b_0 \right) a_1 \\
 & - \left( (\gamma^4 - 6\gamma^2 + 4) a_0 - b_0\gamma^4 + 5\gamma^2 b_0 - 4b_0 \right) b_1 \left. \right) c_0^3 + \left( (-2\gamma^4 a_0 + 4\gamma^2 b_0 - 4b_0) c_1^4 \right. \\
 & + \left( \left( (-3\gamma^4 + 2\gamma^2 - 4) a_0 + 8\gamma^2 b_0 - 8b_0 \right) a_1^2 + 4 \left( (\gamma^4 - \gamma^2 + 1) a_0 - 1/2b_0\gamma^4 - 1/2\gamma^2 b_0 + b_0 \right) b_1 a_1 \right. \\
 & \left. + 2 \left( (\gamma^4 + 10\gamma^2 - 10) a_0 - 2b_0 (\gamma^4 + 2\gamma^2 - 2) \right) a_0 \right) c_1^2 + \left( 2\gamma^2 a_0 + 4\gamma^2 b_0 - 4a_0 - 4b_0 \right) a_1^4 \\
 & - 2 \left( a_0 (\gamma^2 - 2) + 2\gamma^2 b_0 - 2b_0 \right) b_1 a_1^3 + \left( (8\gamma^2 - 12) a_0^2 - 4a_0 b_0 \gamma^2 \right) a_1^2 \\
 & - 6 \left( (\gamma^2 - 4/3) a_0 - 1/3\gamma^2 b_0 \right) a_0 b_1 a_1 - 8 (\gamma - 1) (\gamma + 1) a_0^2 (-b_0 + a_0) \left. \right) c_0^2 \\
 & + 8c_1 \left( \left( (1/4a_0 \left( (\gamma^4 - 2) a_0 - 2\gamma^2 b_0 \right) a_1 - 1/4 \left( (\gamma^4 + \gamma^2 - 2) a_0 - b_0\gamma^4 \right) a_0 b_1 \right) c_1^2 \right. \right. \\
 & - 1/2a_0 \left( \gamma^2 b_0 + a_0 \right) a_1^3 + 1/2a_0 b_1 \left( \gamma^2 b_0 + a_0 \right) a_1^2 - (\gamma - 1) (\gamma + 1) a_0^2 (-b_0 + a_0) a_1 \\
 & \left. \left. + (\gamma - 1) (\gamma + 1) b_1 a_0^2 (-b_0 + a_0) \right) c_0 - 8a_0 \left( 1/8 \left( \gamma^4 a_0 - 4\gamma^2 b_0 + 4b_0 \right) a_0 c_1^4 \right. \right. \\
 & \left. \left. + \left( 1/2a_0 \left( -\gamma^2 b_0 + a_0 + 2b_0 \right) a_1^2 + 1/4b_1 (\gamma^2 - 2) a_0 (a_0 + b_0) a_1 - (\gamma - 1) (\gamma + 1) a_0^2 (-b_0 + a_0) \right) c_1^2 + 1/2a_0 a_1^3 (a_1 - b_1) (a_0 + b_0) \right) , \\
 N_K(s) := & 4 \left( 1/4\gamma^2 c_1 c_0^4 + \left( (3/2a_1 - 1/2b_1) c_1^2 + \left( (-\gamma^2 + 1) a_1 + 1/2 (\gamma - 1) (\gamma + 1) b_1 \right) a_0 + 1/2 \left( a_1^3 - a_1^2 b_1 + 2b_0 a_1 - b_0 b_1 \right) (\gamma^2 - 1) \right) c_0^3 \right. \\
 & + c_1 \left( 1/2c_1^4 + \left( -1/2a_0 - 1/4 (\gamma^2 - 6) a_1^2 - a_1 b_1 + b_0 \right) c_1^2 + (\gamma - 1) (\gamma + 1) a_0^2 \right. \\
 & + \left( 1/2 (\gamma^2 + 1) a_1^2 - 1/2a_1 b_1 - (\gamma - 1) (\gamma + 1) b_0 \right) a_0 \\
 & - 1/2a_1 \left( (\gamma^2 - 2) a_1^3 + (-\gamma^2 + 2) b_1 a_1^2 + 2b_0 (\gamma^2 - 2) a_1 - b_0 b_1 (\gamma^2 - 2) \right) \left. \right) c_0^2 \\
 & + \left( \left( (1/2\gamma^2 a_0 a_1 + b_0 (a_1 - 1/2b_1)) c_1^4 + \left( (-2\gamma^2 + 2) a_1 + 1/2 (\gamma - 1) (\gamma + 1) b_1 \right) a_0^2 \right. \right. \\
 & + \left( a_1^3 \gamma^2 - b_1 \gamma^2 a_1^2 + 2b_0 (\gamma^2 - 1) a_1 - 1/2b_0 (\gamma - 1) b_1 (\gamma + 1) a_0 + a_1^2 b_0 (-b_1 + a_1) \right) c_1^2 \left. \right) c_0 \\
 & + c_1 a_0 \left( -1/4a_0 (\gamma^2 - 2) c_1^4 + \left( (\gamma - 1) (\gamma + 1) a_0^2 + \left( -1/2 (\gamma^2 - 3) a_1^2 + 1/2b_1 (\gamma^2 - 2) a_1 \right. \right. \right. \\
 & \left. \left. - (\gamma - 1) (\gamma + 1) b_0 \right) a_0 + a_1 b_0 (a_1 - 1/2b_1) \right) c_1^2 + a_1^3 (-b_1 + a_1) (a_0 + b_0) \left. \right) \gamma^2 s \\
 & + 4 \left( -1/4\gamma^2 c_0^5 + 1/4c_1 \left( 2\gamma^2 a_1 - 8a_1 + 2b_1 \right) c_0^4 + \left( -1/2c_1^4 + \left( 1/2a_0 + \left( -1/4\gamma^2 - 1 \right) a_1^2 + a_1 b_1 - b_0 \right) c_1^2 + \left( -\gamma^2 + 1 \right) a_0^2 \right. \right. \\
 & + \left( (-3/2 + 1/2\gamma^2) a_1^2 + \left( -1/2\gamma^2 b_1 + b_1 \right) a_1 + (\gamma - 1) (\gamma + 1) b_0 \right) a_0 \\
 & - 1/2 \left( a_1^3 - a_1^2 b_1 + 2b_0 a_1 - b_0 b_1 \right) a_1 \left. \right) c_0^3 + c_1 \left( (-1/2a_0 a_1 - b_0 a_1 + 1/2b_0 b_1) c_1^2 \right. \\
 & + (\gamma - 1) (\gamma + 1) (a_1 + 1/2b_1) a_0^2 + \left( (-1/2\gamma^2 - 1/2) a_1^3 + 1/2 (\gamma^2 + 1) b_1 a_1^2 \right. \\
 & + \left( -\gamma^2 + 1 \right) b_0 a_1 - 1/2b_0 (\gamma - 1) b_1 (\gamma + 1) a_0 - 1/2a_1 \left( 2a_1^2 b_0 - 2a_1 b_0 b_1 \right) \left. \right) c_0^2 \\
 & + \left( \left( (1/4\gamma^2 - 1/2) a_0^2 c_1^4 + \left( (-\gamma^2 + 1) a_0^3 + \left( (-3/2 + 1/2\gamma^2) a_1^2 + \left( 3/2b_1 - \gamma^2 b_1 \right) a_1 \right. \right. \right. \right. \right. \\
 & \left. \left. + (\gamma - 1) (\gamma + 1) b_0 \right) a_0^2 + \left( -a_1^2 b_0 + 1/2a_1 b_1 b_0 \gamma^2 \right) a_0 \right) c_1^2 - a_0 a_1^3 (-b_1 + a_1) (a_0 + b_0) \left. \right) c_0 \\
 & + c_1^3 a_0 \left( 1/2 (\gamma - 1) (\gamma + 1) b_1 a_0^2 - 1/2b_0 (\gamma - 1) b_1 (\gamma + 1) a_0 \right) \gamma^2 .
 \end{aligned}$$

Un logiciel de calcul symbolique tel que *Maple* permet de vérifier que pour  $\gamma = \gamma_{\text{opt}}$ , nous obtenons bien une réduction d'ordre du correcteur, le coefficient en  $s^2$  du dénominateur de  $K_\gamma$  s'annulant pour  $\gamma = \gamma_{\text{opt}}$ .

Il est évident, à la vue de la taille de  $K_\gamma$  pour  $n = 2$ , que celui-ci risque d'être très imposant pour  $n = 3$ , voire  $n = 4$ . Ceux-ci sont cependant calculables. Nous choisissons de ne pas les afficher ici, leurs expressions symboliques étant volumineuses, mais facilement stockables dans un logiciel de calcul formel tel que *Maple*.

En pratique, nous préférons l'utilisation de correcteurs légèrement sous-optimaux, en choisissant  $\gamma > \gamma_{\text{opt}}$ . En effet, l'implémentation du calcul de  $\gamma_{\text{opt}}$  est très simple puisque nous disposons de formules explicites de celui-ci en fonction des paramètres de  $G$ . Il suffit alors d'évaluer les expressions du Corollaire 2.1 pour  $\gamma > \gamma_{\text{opt}}$ . Une telle opération est réalisée pour un faible coût de calcul contrairement à une  $\gamma$ -itération par exemple, ou à n'importe quel algorithme d'optimisation.

### 4.1.3 Exemples

Dans cette section, nous appliquons les formules explicites trouvées précédemment à des systèmes présentant un faible nombre de paramètres afin de montrer l'intérêt de nos méthodes.

#### Système d'ordre 2 : masse-ressort avec extrémité encastrée

Soit un système masse-ressort illustré par la Figure 4.2 et, notamment, utilisé dans [ZDG96, §10.2]. Une masse  $m$  est reliée à un support fixe par un ressort de raideur  $k$  et un amortisseur d'amplitude  $\xi$ . Nous étudions le déplacement de la masse  $m$ , noté  $y_1$ , lorsque  $m$  est excité par une force  $e_1$ .

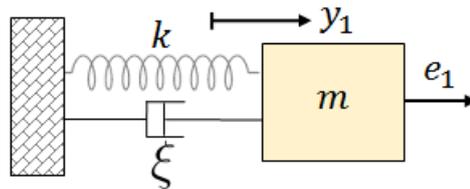


FIGURE 4.2 – Système masse-ressort encastré

La fonction de transfert du système à étudier est alors donnée par :

$$G := \frac{y_1}{e_1} = \frac{c_0}{s^2 + a_1 s + a_0}, \quad c_0 := \frac{1}{m}, \quad a_1 := \frac{\xi}{m}, \quad a_0 := \frac{k}{m}.$$

Il s'agit d'un système d'ordre 2. De plus, comme  $c_1 = 0$ , alors les expressions sous formes closes sont un peu plus simples que dans le cas général. De plus, nous pouvons noter que  $\gamma_{\text{opt}}$  dépend uniquement des paramètres  $\tau^{-1} := G(0) = c_0/a_0 = 1/k$  et  $\rho := a_1/\sqrt{a_0} = \xi/\sqrt{k m}$ . Un tracé de  $\gamma_{\text{opt}}(\tau, \rho)$  est donné dans la Figure 4.3 en utilisant les formules explicites obtenues.

Par exemple, pour  $m = 1$ ,  $k = 1$ ,  $\xi \in \{0, 1/2, 1\}$ , les marges de stabilité garanties et réelles sont données dans la Table 4.6. Rappelons que l'on note  $M_G(G, K)$  (resp.  $M_\Phi(G, K)$ ) la marge de gain (resp. de phase) de la boucle ouverte  $KG$ , et que nous définissons les marges de stabilité garanties par (voir Remarque 2.6 du Chapitre 2) :

$$\begin{cases} M_G(K, G) \geq m_G := \frac{1 + \gamma^{-1}}{1 - \gamma^{-1}}, \\ M_\Phi(K, G) \geq m_\Phi := 2 \arcsin(\gamma^{-1}). \end{cases} \quad (4.7)$$

Noter que ces valeurs sont obtenues par évaluation d'expressions symboliques et ne nécessite pas de recalculer numériquement un correcteur robuste pour les valeurs différentes de l'amortissement  $\xi$ . Les diagrammes de Bode des correcteurs optimaux sont donnés dans la Figure 4.4, et les diagrammes de Black-Nichols de la boucle ouverte pour des correcteurs optimaux sont données dans la Figure 4.5. Sur ces diagrammes, nous pouvons vérifier la cohérence des marges de gain et de phase obtenues suivant les valeurs de  $\gamma_{\text{opt}}$  attendues.

$\xi$	0	1/2	1
$\gamma_{\text{opt}}$	$\simeq 1.80$	$\simeq 1.37$	$\simeq 1.22$
$m_\Phi(\gamma_{\text{opt}})$ ( $^\circ$ )	$\simeq 67.5$	$\simeq 93.9$	$\simeq 110$
$M_\Phi(G, K_{\gamma_{\text{opt}}})$ ( $^\circ$ )	$\simeq 68.9$	$\infty$	$\infty$
$m_G(\gamma_{\text{opt}})$ (dB)	$\simeq 10.9$	$\simeq 16.1$	$\simeq 20.2$
$M_G(G, K_{\gamma_{\text{opt}}})$ (dB)	$\simeq 99.1$	$\simeq 96.7$	$\simeq 95.7$

TABLE 4.6 –  $\gamma_{\text{opt}}$ , marges de phase  $m_\Phi$  et de gain  $m_G$  garanties (bornes inférieures), et marges de phase  $M_\Phi$  et de gain  $M_G$  réelles pour  $\xi \in \{0, 1/2, 1\}$

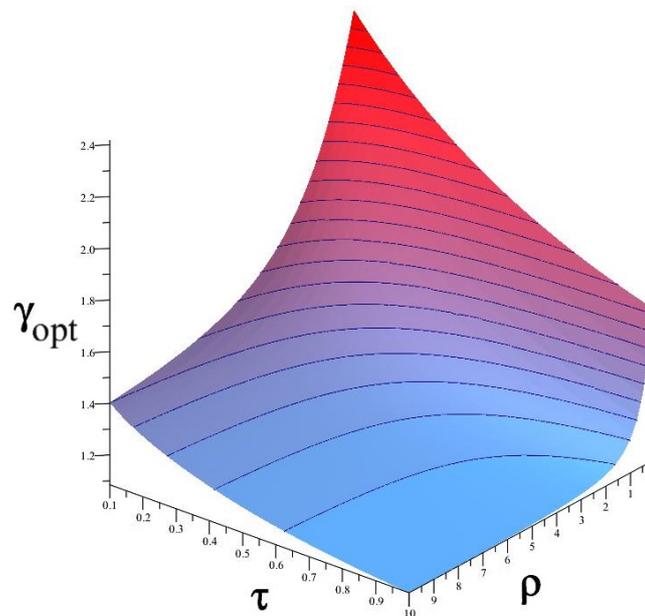


FIGURE 4.3 –  $\gamma_{\text{opt}}$  du système masse-ressort encastré

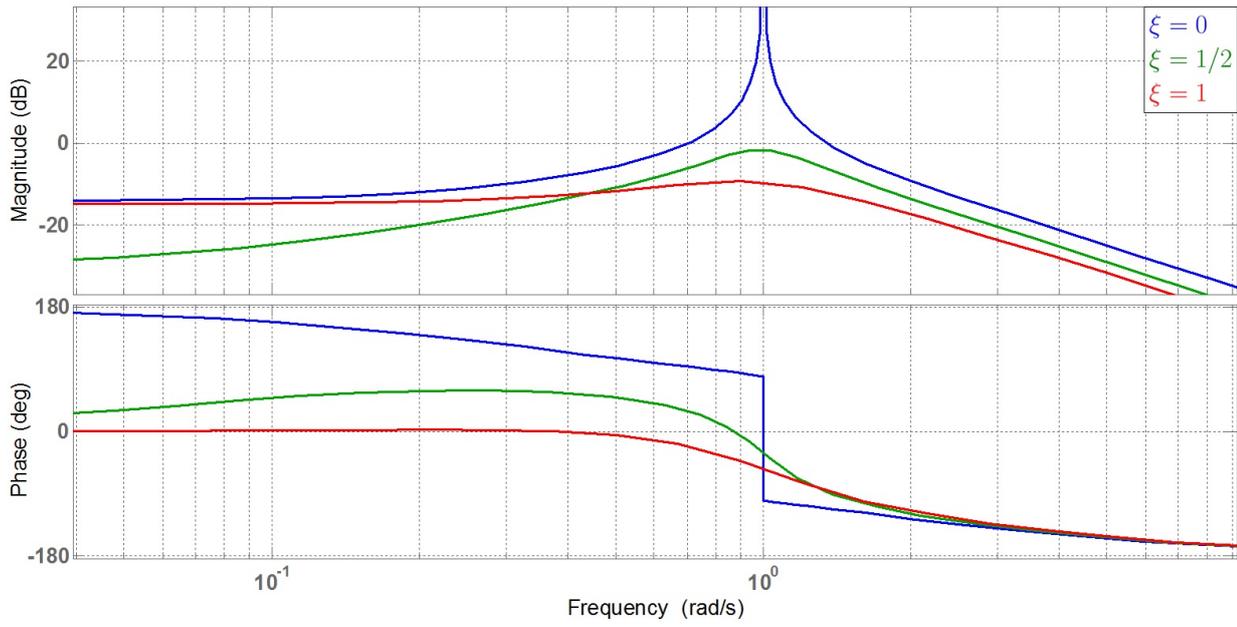


FIGURE 4.4 – Diagramme de Bode de  $K_{\gamma_{\text{opt}}}$  pour  $m = 1$  et  $k = 1$

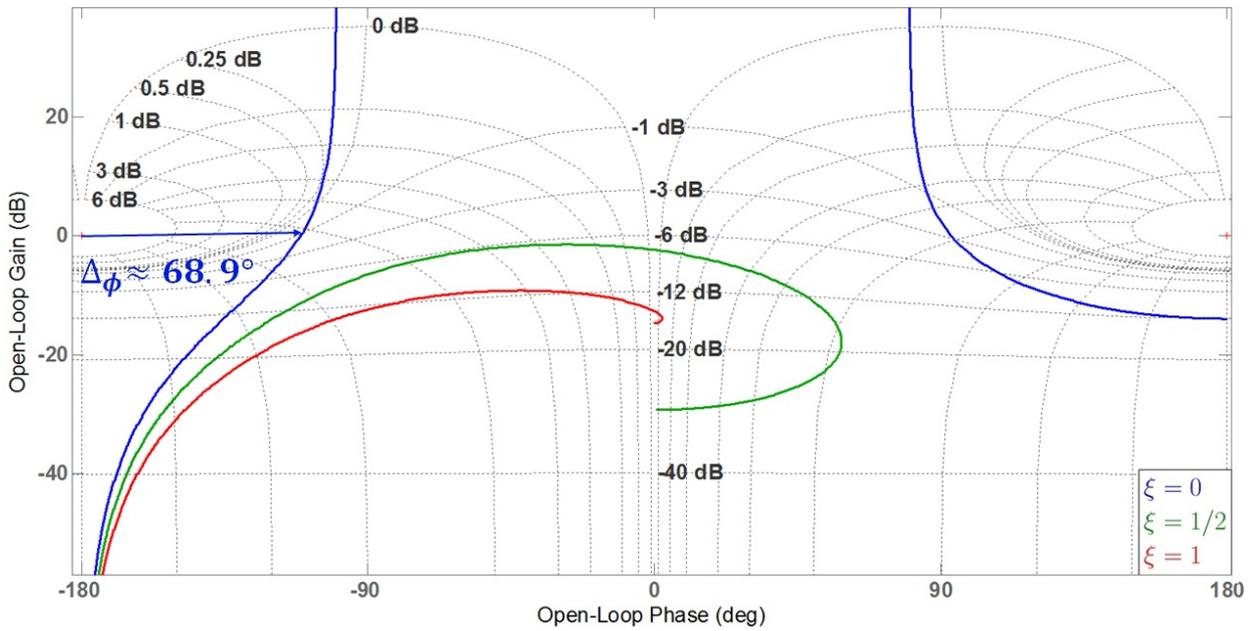


FIGURE 4.5 – Diagramme de Black-Nichols de  $K_{\gamma_{\text{opt}}} G$  pour  $m = 1$  et  $k = 1$

### Système d'ordre 3 : modèle double-poutre d'un viseur gyrostabilité en 2D

En deux dimensions, nous avons vu qu'un viseur pouvait se modéliser par le schéma de la Figure 4.6, à savoir deux poutres de longueur  $l$  et d'inerties  $J_1$  et  $J_2$  reliées par deux ressorts de même raideur  $k$ . Dans cet exemple, nous supposons que les ressorts ne sont pas amortis. La première poutre est entraînée par un moteur de couple  $e_1$ , ce qui constituera l'entrée de notre système, et l'on cherche à stabiliser la vitesse de la deuxième notée  $y_1 := \dot{\theta}$ .

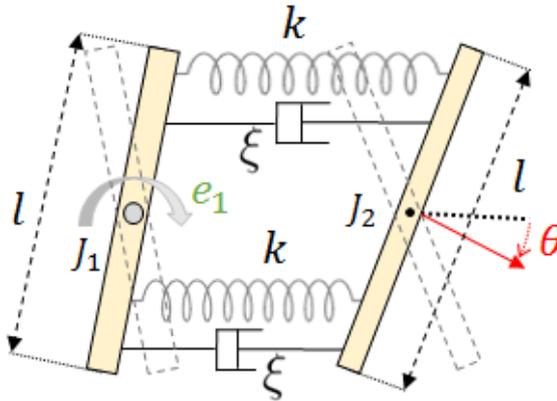


FIGURE 4.6 – Modèle double-poutre d'un viseur gyrostabilisé

La fonction de transfert entre le couple  $e_1$  et la vitesse de rotation  $y_1$  est définie par (voir Annexe F) :

$$G := \frac{y_1}{e_1} = G_0 \frac{a_1 + a_2 s}{s(a_1 + a_2 s + s^2)},$$

$$G_0 := \frac{1}{J_1 + J_2}, \quad a_1 := \frac{2kl^2(J_1 + J_2)}{J_1 J_2}, \quad a_2 := \frac{\xi}{k} a_1.$$

$G$  est un système d'ordre 3 plus simple que le cas général d'un système d'ordre 3 puisque  $c_2 = a_0 = 0$  et  $c_0 = G_0 a_1$ ,  $c_1 = G_0 a_2$ . En reprenant les expressions symboliques d'un système d'ordre 3, et en les simplifiant à l'aide de ces remarques, nous remarquons que  $\gamma_{\text{opt}}$  ne dépend que de deux variables  $\rho := a_2/\sqrt{a_1}$  et  $\tau := G_0/\sqrt{a_1}$  comme on peut le voir sur le tracé de  $\gamma_{\text{opt}}$  de la Figure 4.7.

Nous pouvons aussi reprendre les formules explicites des correcteurs  $H_\infty$  obtenus dans le cas général. Par exemple, pour  $J_1 = J_2 = 1$ ,  $k = 1$ ,  $l = 1$ ,  $\xi \in \{0, 1/2, 1\}$ , les diagrammes de Bode des correcteurs obtenus sont donnés dans la Figure 4.8 et les diagrammes de Black-Nichols de la boucle ouverte sont affichés dans la Figure 4.9. Nous pouvons aussi vérifier que les correcteurs obtenus assurent bien les marges de stabilité garanties par  $\gamma_{\text{opt}}$ , notées  $m_\Phi$  et  $m_G$ , comme l'indique la Table 4.7.

$\xi$	0	1/2	1
$\gamma_{\text{opt}}$	$\simeq 1.86$	$\simeq 1.48$	$\simeq 1.46$
$m_\Phi(\gamma_{\text{opt}})$ ( $^\circ$ )	$\simeq 64.7$	$\simeq 84.7$	$\simeq 86.2$
$M_\Phi(G, K_{\gamma_{\text{opt}}})$ ( $^\circ$ )	$\simeq 73.5$	$\simeq 86.5$	$\simeq 86.6$
$m_G(\gamma_{\text{opt}})$ (dB)	$\simeq 10.37$	$\simeq 14.2$	$\simeq 14.5$
$M_G(G, K_{\gamma_{\text{opt}}})$ (dB)	$\simeq 10.5$	$\simeq 91$	$\simeq 96.4$

TABLE 4.7 –  $\gamma_{\text{opt}}$ , marges de phase  $m_\Phi$  et de gain  $m_G$  garanties, et marges de phase  $M_\Phi$  et de gain  $M_G$  réelles pour  $\xi \in \{0, 1/2, 1\}$

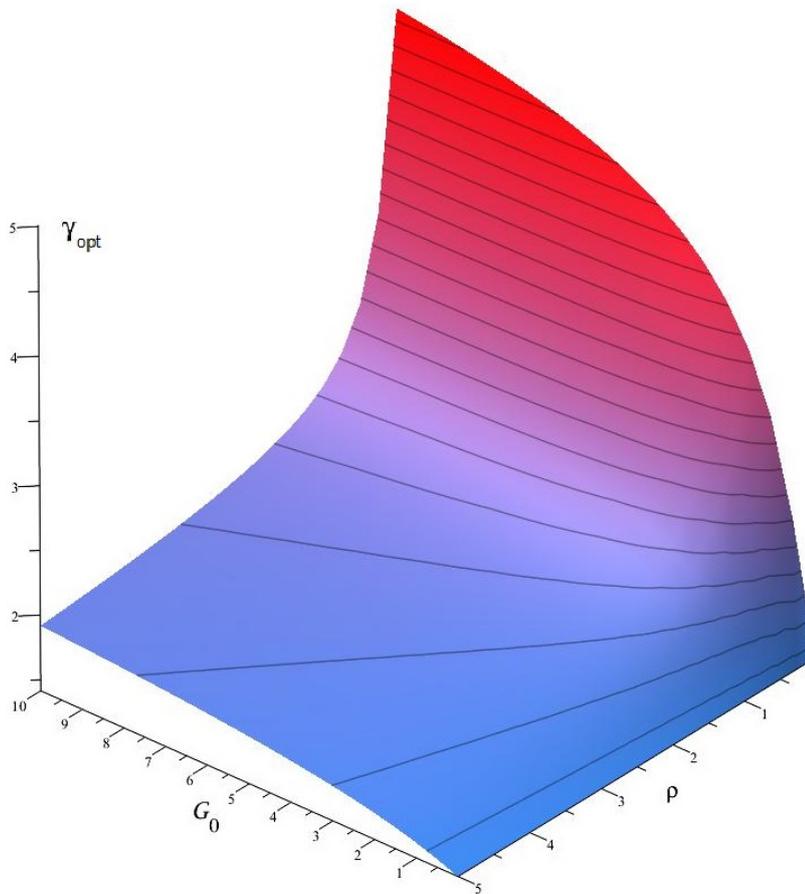


FIGURE 4.7 –  $\gamma_{\text{opt}}$  pour un système double poutre

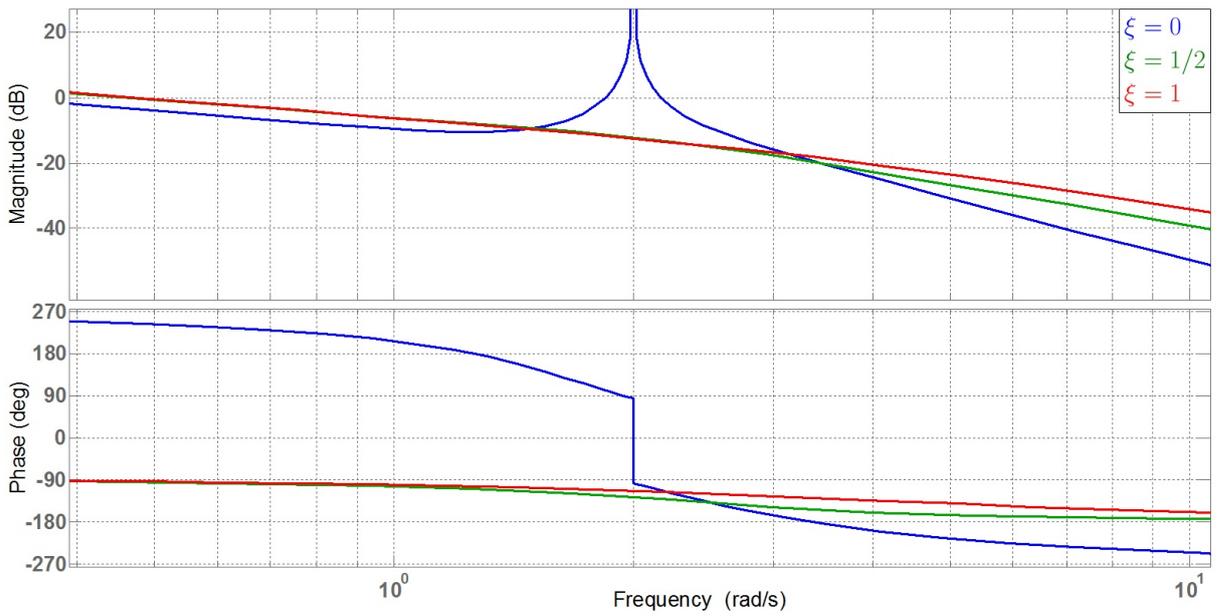


FIGURE 4.8 – Diagramme de Bode de  $K_{\gamma_{\text{opt}}}$  pour  $J_1 = J_2 = 1$ ,  $k = 1$  et  $l = 1$

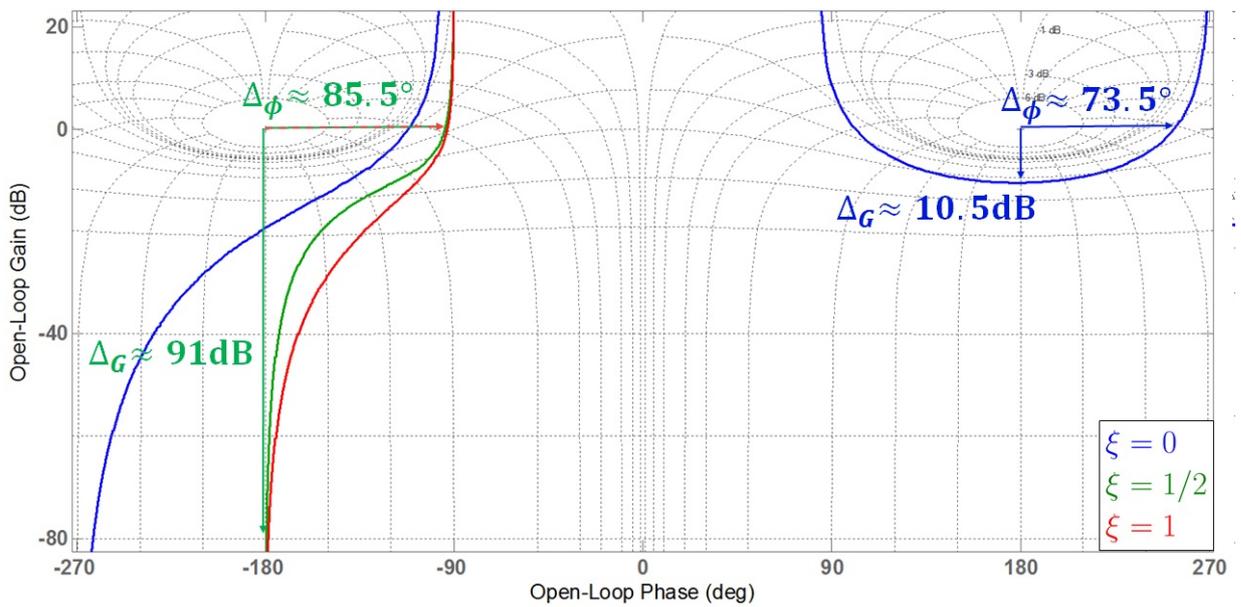


FIGURE 4.9 – Diagramme de Black-Nichols de  $K_{\gamma_{\text{opt}}} G$  pour  $J_1 = J_2 = 1$ ,  $k = 1$  et  $l = 1$

**Système d'ordre 4 : système masse-ressort libre**

Enfin, nous choisissons de revenir sur un système masse-ressort pour illustrer nos démarches dans le cadre des systèmes d'ordre 4. Ce système est un exemple standard dans la théorie du contrôle et constitue un benchmark classique. Voir, par exemple, [WB92, Vin00, ACA<sup>+</sup>99]. Il s'agit d'un modèle de deux masses  $m_1$  et  $m_2$  reliées par un ressort de raideur  $k$  et d'amortissement  $f$ . Celui-ci est représenté dans la Figure 4.10.

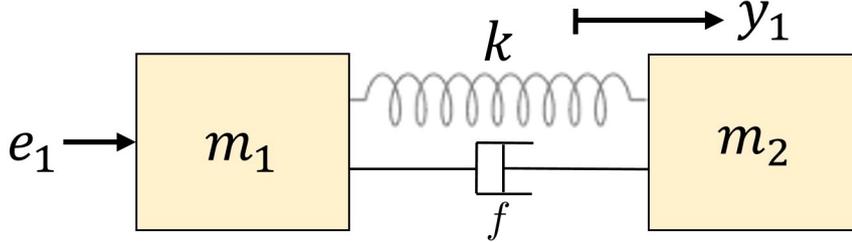


FIGURE 4.10 – Modèle deux-masses-ressort

Avec les notations de la Figure 4.10, nous étudions le déplacement de la deuxième masse, notée  $y_1$ , lorsque la première est soumise à une force  $e_1$  l'entraînant en translation. Nous considérons alors le système  $G$  entre l'entrée  $e_1$  et la sortie  $y_1$  défini par :

$$\frac{y_1}{e_1} := G = \frac{a_3 s + a_2}{m s^2 (s^2 + a_3 s + a_2)},$$

$$m := m_1 + m_2 > 0, \quad a_2 := \frac{(m_1 + m_2) k}{m_1 m_2} > 0, \quad a_3 := \frac{(m_1 + m_2) f}{m_1 m_2} > 0.$$

Au travers de cet exemple, nous proposons de montrer que le nombre de paramètres influe grandement sur la complexité des calculs sous-jacents. Nous distinguerons deux cas : le cas où  $f = 0$ , réduisant de 1 le nombre de paramètres, et le cas où  $f \neq 0$ . Considérer un système non-amorti permet de traiter *le pire des cas* puisqu'un ressort peut avoir un amortissement faible mais non nul. Un correcteur stabilisant le système sans amortissement stabilisera le même système avec amortissement (en effet, sur un diagramme de Black-Nichols, nous remarquons qu'un correcteur stabilisant un système sans amortissement "place le mode mécanique entre les deux points critiques"). C'est donc un bon point de départ. Détaillons la démarche pour  $f = 0$ .

- Étude du système sans amortissement :

La fonction de transfert a la forme suivante :

$$G = \frac{y_1}{e_1} = \frac{c_0}{s^2 (s^2 + a_2)}, \quad a_2 > 0, \quad c_0 := \frac{a_2}{m} > 0.$$

Dans cet exemple, le problème ne fait intervenir que 2 paramètres :

$$\theta_0 := (a_2, c_0).$$

De plus, le système  $\mathcal{B}$  s'écrit alors

$$\begin{cases} \mathcal{B}_0 := b_0^2 - d_0 = 0, \\ \mathcal{B}_1 := b_1^2 - 2 b_0 b_2 - d_2 = 0, \\ \mathcal{B}_2 := b_2^2 - 2 b_1 b_3 + 2 b_0 - d_4 = 0, \\ \mathcal{B}_3 := b_3^2 - 2 b_2 - d_6 = 0, \end{cases} \quad (4.8)$$

où les paramètres  $d_{2i}$  sont définis par :

$$d_0 := c_0^2, \quad d_2 := 0, \quad d_4 := a_2^2, \quad d_6 := -2a_2. \quad (4.9)$$

En utilisant  $\mathcal{B}_0 = 0$ , nous trouvons que  $X > 0$  satisfait  $b_0 = c_0 > 0$  (voir Proposition 3.11). En effet, nous supposons  $c_0 \neq 0$  car  $c_0 = 0$  donne  $G = 0$ , cas que nous excluons. Nous pouvons alors retirer cette équation en substituant  $b_0$  par  $c_0$  dans (4.8) et ne travailler que sur le reste du système  $\mathcal{B}$ , noté  $\tilde{\mathcal{B}}$ . Une paramétrisation de  $\tilde{\mathcal{B}}$  se calcule facilement à la main :

$$\tilde{\mathcal{R}} := \begin{cases} b_1 = \frac{b_3^4 + 4a_2b_3^2 + 8c_0}{8b_3}, \\ b_2 = \frac{1}{2}b_3^2 + a_2, \\ b_3 = b_3, \\ \tilde{p}(b_3) := b_3^8 + 8a_2b_3^6 + 16(a_2^2 - 3c_0)b_3^4 - 64a_2c_0b_3^2 + 64c_0^2 = 0. \end{cases} \quad (4.10)$$

En appliquant le changement de variable  $\tau = (8c_0)^{-1/2}b_3^2$  ( $c_0 > 0$  par hypothèse) au polynôme  $\tilde{p}$ , nous obtenons :

$$\tilde{p}_\tau := c_0\tau^4 + \sqrt{8c_0}a_2\tau^3 + 2(a_2^2 - 3c_0)\tau^2 - \sqrt{8c_0}a_2\tau + c_0 = 0. \quad (4.11)$$

Remarquons qu'en écrivant  $\tilde{p}_\tau$  sous la forme

$$\tilde{p}_\tau = \sum_{i=0}^4 \alpha_i \tau^i,$$

alors  $\alpha_i = (-1)^i \alpha_{4-i}$  pour  $i = 0, 1, \dots, 4$ . Nous pouvons alors appliquer le changement de variable  $\varsigma := \tau - \tau^{-1}$  afin d'obtenir :

$$\tilde{p}_\varsigma := \varsigma^2 + 2\sqrt{\frac{2a_2^2}{c_0}}\varsigma + 2\frac{a_2^2}{c_0} - 4 = 0. \quad (4.12)$$

La résolution de l'équation polynomiale  $\tilde{p}(b_3) = 0$  de degré 8 est alors réduite à la résolution de  $\tilde{p}_\varsigma(\varsigma) = 0$  de degré 2. Le discriminant de  $\tilde{p}_\varsigma$  en  $\varsigma$  est  $16 > 0$  et donc  $\tilde{p}_\varsigma$  a deux racines réelles distinctes en  $\varsigma$  :

$$\begin{cases} \varsigma_+ := -\sqrt{\frac{2a_2^2}{c_0}} + 2, \\ \varsigma_- := -\sqrt{\frac{2a_2^2}{c_0}} - 2, \end{cases} \quad \varsigma_- < 0, \quad \varsigma_+ > \varsigma_-, \quad |\varsigma_-| > |\varsigma_+|.$$

Alors,  $\tau$  vérifie l'équation  $\tau^2 - \tau\varsigma - 1 = 0$  dont le discriminant en  $\tau$  est  $\varsigma^2 + 4 > 0$ . Ainsi, les solutions (4.11) sont toutes réelles et s'écrivent sous la forme suivante :

$$\begin{cases} \tau_+ := \frac{1}{2} \left( \varsigma_\pm + \sqrt{\varsigma_\pm^2 + 4} \right), \\ \tau_- := \frac{1}{2} \left( \varsigma_\pm - \sqrt{\varsigma_\pm^2 + 4} \right), \end{cases} \quad \varsigma_\pm \in \{\varsigma_+, \varsigma_-\}.$$

Trivialement,  $\tau_+$  est la racine réelle maximale de (4.11), ce qui nous donne directement la racine réelle maximale de  $\tilde{p}$  :

$$\sigma := \max \{b_3 \mid \tilde{p}(b_3) = 0\} = \sqrt{2} \sqrt{(\sqrt{2c_0} - a_2) + \sqrt{(\sqrt{2c_0} - a_2)^2 + 2c_0}}. \quad (4.13)$$

Le Théorème 3.11 permet d'écrire la forme de  $X$  solution de  $\mathcal{R} = 0$  en fonction des  $b_i$  définis en (4.10) :

$$X := \begin{pmatrix} c_0 b_1 & c_0 b_2 & c_0 b_3 & c_0 \\ c_0 b_2 & -c_0 b_3 + b_1 b_2 & b_1 b_3 - c_0 & b_1 \\ c_0 b_3 & b_1 b_3 - c_0 & b_2 b_3 - b_1 & b_2 - a_2 \\ c_0 & b_1 & b_2 - a_2 & b_3 \end{pmatrix}.$$

La matrice  $\mathcal{Q}$  est aussi très simple et s'écrit :

$$\mathcal{Q} := \begin{pmatrix} 0 & 0 & 0 & c_0^{-1} \\ 0 & 0 & c_0^{-1} & 0 \\ 0 & c_0^{-1} & 0 & -a_2 c_0^{-1} \\ c_0^{-1} & 0 & -a_2 c_0^{-1} & 0 \end{pmatrix}.$$

Pour cet exemple, montrons que l'on peut conclure sur l'expression de  $\gamma_{\text{opt}}$  directement en travaillant sur les valeurs propres de la matrice  $\mathcal{Q}X$ . Le polynôme caractéristique de  $\mathcal{Q}X$  multiplié par  $4/c_0^2$  s'écrit

$$h_s(\lambda, b_3) := 4 c_0 \lambda^4 - \alpha(b_3) \lambda^3 + \beta(b_3) \lambda^2 + \alpha(b_3) \lambda + 4 c_0, \quad (4.14)$$

où :

$$\alpha := b_3^4 + 8 c_0, \quad \beta := -2(b_3^4 + 4 c_0). \quad (4.15)$$

Encore une fois, nous remarquons une symétrie dans  $h_s$ . En considérant le changement de variable  $t := \lambda - \lambda^{-1}$  ( $\lambda = 0$  n'étant pas racine de  $h_s$  car  $c_0 \neq 0$ ),  $h_s$  peut se réécrire sous la forme de :

$$h_t(t, b_3) = 4 c_0 t^2 - \alpha t - 2 b_3^4. \quad (4.16)$$

En particulierisant  $b_3$  en  $\sigma$ , les racines de  $h_t(t, \sigma)$  en  $t$  sont alors :

$$\begin{cases} t_+ := \frac{1}{8 c_0} \left( (\sigma^4 + 8 c_0)^2 + \sqrt{(\sigma^4 + 8 c_0)^2 + 32 c_0 \sigma^4} \right), \\ t_- := \frac{1}{8 c_0} \left( (\sigma^4 + 8 c_0)^2 - \sqrt{(\sigma^4 + 8 c_0)^2 + 32 c_0 \sigma^4} \right). \end{cases} \quad (4.17)$$

Nous en déduisons alors les racines de  $h_s$  :

$$\begin{cases} \lambda_+ := \frac{1}{2} \left( t_\pm + \sqrt{t_\pm^2 + 4} \right), \\ \lambda_- := \frac{1}{2} \left( t_\pm - \sqrt{t_\pm^2 + 4} \right), \end{cases} \quad t_\pm \in \{t_+, t_-\}.$$

Il nous reste alors à déterminer quelle racine de  $h_s(\lambda, \sigma)$  est de plus grande valeur absolue. Notons que  $\sigma > 0$  et que nous avons  $t_+ t_- = -\frac{2\sigma^4}{4c_0} < 0$  et  $t_+ > t_-$ . Ainsi  $t_-$  est nécessairement négative tandis que  $t_+$  est nécessairement positive. De plus, comme  $t_+ + t_- = \frac{\alpha(\sigma)}{4c_0} > 0$ , alors nous avons  $t_+ = |t_+| > -t_-$ , c'est-à-dire  $|t_+| > |t_-|$ . En conséquence, nous avons :

$$\begin{aligned} \left| t_+ + \sqrt{t_+^2 + 4} \right| &> \left| t_+ - \sqrt{t_+^2 + 4} \right|, \\ \left| t_- - \sqrt{t_-^2 + 4} \right| &> \left| t_- + \sqrt{t_-^2 + 4} \right|. \end{aligned}$$

La racine de  $h_s(\lambda, \sigma)$  de plus grande valeur absolue est alors définie par :

$$\lambda_a := \frac{1}{2} \left( t_+ + \sqrt{t_+^2 + 4} \right). \quad (4.18)$$

Nous en déduisons finalement l'expression explicite de  $\gamma_{\text{opt}}$  suivante :

$$\gamma_{\text{opt}}(\sigma) := \sqrt{1 + \left( \frac{1}{2} \left( t_+ + \sqrt{t_+^2 + 4} \right) \right)^2}.$$

Un tracé de  $\gamma_{\text{opt}}$  est alors donné dans la Figure 4.11.

Remarquons simplement que pour un tel système, *l'expression explicite de  $\gamma_{\text{opt}}$  en fonction des paramètres est très simple.*

Montrons maintenant l'impact de l'ajout d'un paramètre  $f$ .

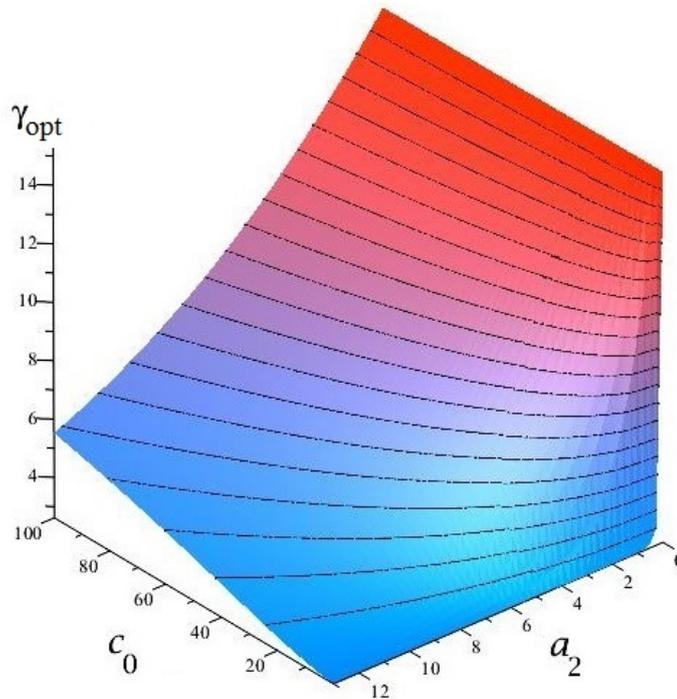


FIGURE 4.11 – Tracé de  $\gamma_{\text{opt}}$  en fonction de  $a_2$  et  $c_0$

• Étude du système avec amortissement :

Dans ce cas, le problème fait intervenir les 3 paramètres suivants :

$$\theta := (a_2, a_3, m).$$

Le système  $\mathcal{B}$  s'écrit toujours

$$\begin{cases} \mathcal{B}_0 := b_0^2 - d_0 = 0, \\ \mathcal{B}_1 := b_1^2 - 2b_0b_2 - d_2 = 0, \\ \mathcal{B}_2 := b_2^2 - 2b_1b_3 + 2b_0 - d_4 = 0, \\ \mathcal{B}_3 := b_3^2 - 2b_2 - d_6 = 0, \end{cases}$$

où les paramètres  $d_{2i}$  s'écrivent alors :

$$d_0 = \frac{a_2^2}{m^2}, \quad d_2 = \frac{a_3^2}{m^2}, \quad d_4 = a_2^2, \quad d_6 = a_3^2 - 2a_2. \quad (4.19)$$

Cet exemple fait intervenir un plus grand nombre de paramètres que dans le cas sans amortissement, mais un plus faible nombre de paramètres que dans le cas général. Notamment, en utilisant  $\mathcal{B}_0 = 0$ , nous trouvons que  $X > 0$  satisfait  $b_0 = a_2/m$  (voir Proposition 3.11). Nous supposons aussi  $a_2 \neq 0$  car  $a_2 = 0$  donne un système sans mode mécanique, cas que nous excluons. Nous pouvons alors retirer cette équation et ne plus travailler que sur  $\tilde{\mathcal{B}}$ . Une paramétrisation de  $\tilde{\mathcal{B}}$  se calcule à la main :

$$\tilde{\mathcal{R}} := \begin{cases} b_1 &= \frac{b_3^4 - 2(a_3^2 - 2a_2)b_3^2 + (a_3^2 - 2a_2)^2 + 8\frac{a_2}{m} - 4a_2^2}{8b_3}, \\ b_2 &= \frac{1}{2}(b_3^2 - a_3^2) + a_2, \\ b_3 &= b_3, \\ \tilde{p}(b_3) &= b_3^8 + 4(2a_2 - a_3^2)b_3^6 + \left(6a_3^4 - 24a_2a_3^2 + 16a_2^2 - 48\frac{a_2}{m}\right)b_3^4 \\ &\quad + \left(24a_2a_3^4 - 4a_3^6 - 32\left(a_2^2 - \frac{a_2}{m} + \frac{2}{m^2}\right)a_3^2 - \frac{64a_2^2}{m}\right)b_3^2 \\ &\quad + \left(4a_2a_3^2 - a_3^4 - 8\frac{a_2}{m}\right)^2. \end{cases}$$

Dans le cas général, nous avons établi que l'expression par radicaux de la racine réelle maximale de  $\tilde{p}$  pouvait varier suivant les valeurs des paramètres. Ici, le nombre de paramètres étant réduit, un calcul de décomposition en cellules à partir de la variété discriminante de  $\tilde{p}$  montre que la racine réelle maximale de  $\tilde{p}$  a une seule et unique expression par radicaux sur la totalité de l'espace des paramètres (excepté en sur les variétés singulières  $\mu_1 = 0$  et  $p_4 = 0$  définies dans la Table 3.6) :

$$\tilde{\sigma} = \sqrt{t(1, 1) + d_6},$$

où  $t(1, 1)$  est défini dans la Table 3.5. Nous obtenons alors  $X(\tilde{\sigma}) > 0$ . En utilisant la Table 4.2, nous déduisons  $\mathcal{Q}$ , puis  $Y := \mathcal{Q}X\mathcal{Q} > 0$ . Nous pouvons alors calculer le polynôme caractéristique de  $YX$ , noté  $h$ , et en déduire sa racine réelle maximale comme effectué dans le cas général (voir Table 4.5). Contrairement au cas sans amortissement, notons qu'ici toutes ces expressions ne sont pas affichables du fait de leur taille plus imposante.

L'utilisation d'un logiciel de calcul formel tel que *Maple* permet de simplifier l'expression de  $\gamma_{\text{opt}}$  et de montrer que celui-ci ne dépend que de deux paramètres suivants qui sont des combinaisons des 3 autres :

$$G_r := \frac{1}{m a_2}, \quad \rho := \frac{a_3}{\sqrt{a_2}}.$$

Nous représentons alors le graphe de  $\gamma_{\text{opt}}$  en fonction de  $G_r$  et  $\rho$  dans la Figure 4.12.

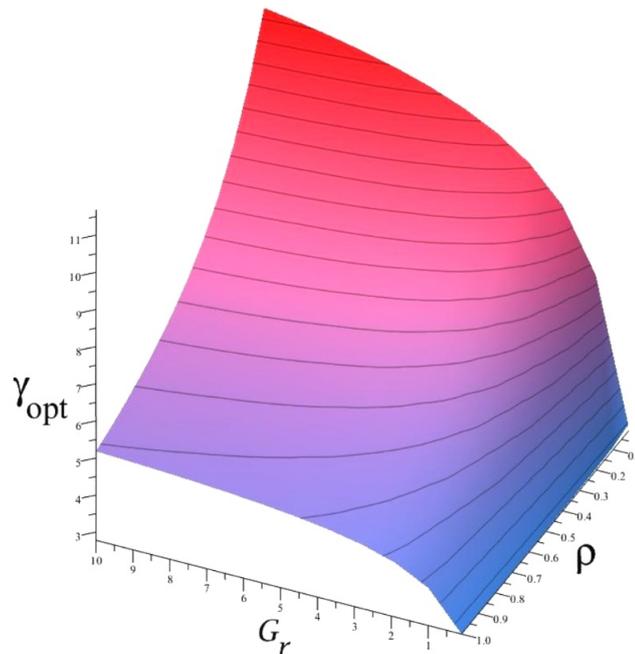


FIGURE 4.12 – Tracé de  $\gamma_{\text{opt}}$  en fonction de  $G_r$  et  $\rho$

• Comparaison des deux cas :

Le lecteur aura facilement remarqué que les expressions se complexifient rapidement ne serait-ce qu'en rajoutant un paramètre supplémentaire. Dans le cas sans amortissement, en n'ayant recours qu'à des expressions par radicaux de polynômes de degré 2, nous déduisons des expressions très simples de  $\gamma_{\text{opt}}$ . Dans le cas avec amortissement, nous devons utiliser des racines par radicaux de polynômes de degré 4 afin de déterminer une expression de  $\gamma_{\text{opt}}$ . Cette dernière est alors bien plus complexe que dans le cas sans amortissement. De plus, le cas sans amortissement se traite assez aisément à la main, alors que le cas avec amortissement nécessite le calcul d'une variété discriminante et d'une décomposition de l'espace des paramètres en cellules, choses impossibles à faire à la main.

## 4.2 Résolution symbolique/numérique du problème de commande robuste

Pour des systèmes d'ordre  $n \geq 5$ , nous ne pouvons plus généralement trouver de formule par radicaux pour  $\gamma_{\text{opt}}$ . En effet, comme nous l'expliquions à la Section 4.1, l'ordre des polynômes  $p$  et  $h$  entrant en jeu est trop élevé et nous sommes limités par la théorie de Galois [Mut80]. Une étape numérique est alors nécessaire. La question est maintenant de savoir quand évaluer numériquement les valeurs des paramètres.

Soit  $n \geq 5$ . Supposons que les coefficients  $a$  et  $c$  soient fixés à des valeurs rationnelles. De plus, supposons que l'on dispose d'une paramétrisation rationnelle des solutions du système  $\mathcal{B}$  suivant un polynôme univarié  $p$  en  $b_{n-1}$  (voir Section 3.5.3). Nous disposons alors d'une paramétrisation rationnelle de  $X(b_{n-1})$  solution de  $\mathcal{R} = 0$ . Considérons alors le polynôme caractéristique de  $\mathcal{Q}X$  défini par

$$h_\star(\lambda, b_{n-1}) = 0,$$

ainsi que le système  $\mathcal{T}_\star$  suivant :

$$\mathcal{T}_\star := \begin{cases} p(b_{n-1}) = 0, \\ h_\star(\lambda, b_{n-1}) = 0. \end{cases}$$

Nous disposons alors d'un système polynomial triangulaire en deux variables,  $b_{n-1}$  et  $\lambda$ , dont les coefficients sont rationnels.

Rappelons que l'objectif est de trouver la plus grande valeur absolue  $\lambda_a$  des racine réelles de  $h_\star$  (polynôme en  $\lambda$ ) sachant que  $\sigma$  est la racine réelle maximale de  $p$  (polynôme en  $b_{n-1}$ ). Pour résoudre ce système, une première idée serait d'abord de calculer les racines de  $p$  en  $b_{n-1}$  à l'aide d'un algorithme numérique (citons par exemple la *méthode de Newton* ou la *méthode des trapèzes* [Nou85] pour les plus célèbres), puis d'en choisir la racine réelle maximale  $\sigma$ . En injectant la valeur numérique obtenue pour  $\sigma$  dans  $h_\star$ , nous obtenons alors un polynôme univarié en  $\lambda$  à coefficients numériques, dont nous pouvons calculer les racines de la même manière que précédemment.

Une telle méthode a l'inconvénient de ne pas être *certifiée* [CA76, RZ04], plus particulièrement lorsque  $h_\star$  n'est pas sans facteur carré en  $\lambda$ . Pour expliquer ceci, supposons que  $h_\star$  n'est pas sans facteur carré. La substitution d'une approximation numérique d'une racine de  $p$  dans  $h_\star$  donne un polynôme  $h'_\star$ , potentiellement sans facteur carré. L'isolation des racines de  $h'_\star$  peut alors être erronée : elle peut par exemple retourner des valeurs complexes, alors que  $h_\star$  ne possède que des racines réelles. Afin d'obtenir un calcul certifié des racines, nous pouvons utiliser les propriétés du système polynomial triangulaire  $\mathcal{T}_\star$  afin de calculer une *approximation numérique certifiée de ses solutions* [CA76, RZ04]. Une approximation certifiée des solutions d'un système d'équations est une représentation des solutions de ce système sous forme d'intervalles contenant les solutions réelles du système. Considérons par exemple l'équation  $p(b_{n-1}) = 0$ . Une approximation certifiée des racines de  $p$  donnera des intervalles pour les valeurs de  $b_{n-1}$  contenant à *coup sûr* les racines de  $p$ . Pour plus de détails sur l'analyse par intervalles, nous référons, par exemple, à [Jau01]. Pour des applications de l'analyse par intervalles, citons, par exemple, [DLDJ<sup>+</sup>07].

Il existe plusieurs manières de procéder pour résoudre  $\mathcal{T}_\star$  de manière certifiée. Nous pouvons mentionner l'algorithme de [CGY07] qui utilise des *bornes séparantes* afin d'isoler les

racines du polynôme  $h_*$  au dessus d'une racine de  $p$ . Nous pouvons aussi mentionner l'algorithme donné dans [BCLM14] qui requiert la décomposition de  $\mathcal{T}_*$  en un ensemble de systèmes triangulaires sans facteurs carrés, de telle sorte que pour chaque racine de  $p$ , le polynôme  $h_*$  soit sans facteur carré.

Nous privilégierons ici de réutiliser l'idée développée dans le Chapitre 3, à savoir calculer une RUR du système  $\mathcal{T}_*$ . Pour un *élément séparant* bien choisi  $\tau := u_1 b_{n-1} + u_2 \lambda$ , les solutions de  $\mathcal{T}_*$  peuvent s'écrire de la manière suivante :

$$\begin{cases} v(\tau) = 0, \\ b_{n-1} = \frac{v_1(\tau)}{w(\tau)}, \\ \lambda = \frac{v_2(\tau)}{w(\tau)}. \end{cases} \quad (4.20)$$

Ensuite, afin de déterminer une solution de (4.20), il est suffisant d'isoler les racines de  $v$  par intervalles, puis d'évaluer les intervalles obtenus dans les expressions  $\frac{v_1(\tau)}{w(\tau)}$  et  $\frac{v_2(\tau)}{w(\tau)}$ . Nous obtenons alors des boîtes isolant les solutions  $(b_{n-1}, \lambda)$  de (4.20). Il suffit alors de choisir celle qui satisfait la condition voulue, c'est-à-dire celle dont la coordonnée  $\lambda$  est de plus grande valeur absolue sachant que la coordonnée  $b_{n-1}$  est maximale.

**Exemple 4.1.** Reprenons l'exemple de la double poutre asservie en vitesse utilisé à la section précédente. Le système est de la forme suivante :

$$G := \frac{y_1}{e_1} = G_0 \frac{a_1 + a_2 s}{s(a_1 + a_2 s + s^2)},$$

$$G_0 := \frac{1}{J_1 + J_2}, \quad a_1 := \frac{2kl^2(J_1 + J_2)}{J_1 J_2}, \quad a_2 := \frac{\xi}{k} a_1.$$

Choisissons les valeurs de paramètres suivantes :

$$J_1 = J_2 = 1, \quad k = 1, \quad l = 1, \quad \xi = \frac{1}{2}.$$

Nous obtenons alors le système polynomial  $\mathcal{T}_*$  suivant :

$$\begin{cases} 4b_2^4 - 1984b_2^2 - 64_2 + 244736 = 0, \\ 256\lambda^3 + (16b_2^4 - 64b_2^3 + 96b_2^2 + 256b_2 - 1152)\lambda^2 \\ + (-8b_2^6 + 32b_2^5 - 96b_2^4 + 512b_2^3 - 672b_2^2 + 1408b_2 - 4864)\lambda \\ + b_2^9 - 2b_2^8 + 16b_2^7 - 56b_2^6 + 28b_2^5 - 224b_2^4 - 352b_2^3 + 1728b_2^2 - 1088b_2 + 6880 = 0. \end{cases} \quad (4.21)$$

Notons que le premier polynôme de (4.21) n'est pas  $p(b_{n-1})$  mais  $\tilde{p}(b_{n-1}, b_0)$  auquel on a particularisé la valeur de  $b_0$ . En effet, pour des systèmes présentant un intégrateur, la valeur de  $b_0$  est triviale. Dans le cas présent, les valeurs de paramètres sont particularisées à des valeurs numériques et nous pouvons travailler sur  $\tilde{p}$  sans perte de généralité.

Le calcul d'une RUR de (4.21) pour l'élément séparant  $t = \lambda$  donne :

$$\left\{ \begin{array}{l} 4t^{12} - 60t^{11} + 1417t^{10} - 1620t^9 - 76048t^8 - 920t^7 + 153366t^6 + 2712t^5 \\ \quad - 78096t^4 + 2004t^3 + 393t^2 - 68t + 4 = 0, \\ \lambda = t, \\ \quad 244t^{10} - 254t^9 - 12860t^8 - 152536t^7 - 53920t^6 + 305036t^5 \\ \quad - 11056t^4 - 161624t^3 + 29580t^2 - 318t - 180 \\ b_2 = \frac{\quad}{24t^{11} - 330t^{10} + 7085t^9 - 7290t^8 - 304192t^7 - 3220t^6 + 460098t^5} \\ \quad + 6780t^4 - 156192t^3 + 3006t^2 + 393t - 34. \end{array} \right. \quad (4.22)$$

L'extraction des racines du polynôme univarié de la RUR donne les intervalles certifiés suivants pour  $t$  (nous utilisons la commande *RootFinding/Isolate* de *Maple*) :

$$\begin{aligned} t &\in \left[ -\frac{825188652418145181}{144115188075855872}, -\frac{206297163104536295}{36028797018963968} \right], \\ t &\in \left[ -\frac{134385842784865663}{1152921504606846976}, -\frac{67192921392432831}{576460752303423488} \right], \\ t &\in \left[ \frac{80711476207981101}{1152921504606846976}, \frac{40355738103990551}{576460752303423488} \right], \\ t &\in \left[ \frac{128609005596969279}{144115188075855872}, \frac{2009515712452645}{2251799813685248} \right], \\ t &\in \left[ \frac{4939673728209055}{4503599627370496}, \frac{158069559302689761}{144115188075855872} \right], \\ t &\in \left[ \frac{1235657290510629919}{144115188075855872}, \frac{38614290328457185}{4503599627370496} \right]. \end{aligned}$$

Nous obtenons alors les boîtes suivantes pour les solutions de (4.21) :

$$\begin{aligned} \lambda &= \left[ \frac{-422496590038090332259}{73786976294838206464}, \frac{-211248295019045166117}{36893488147419103232} \right] & b_2 &= \left[ \frac{-468997241444025982405}{295147905179352825856}, \frac{-234498620722012991195}{147573952589676412928} \right], \\ \lambda &= \left[ \frac{-137611103011702438891}{1180591620717411303424}, \frac{-17201387876462804861}{147573952589676412928} \right], & b_2 &= \left[ \frac{93464837667499525323}{36893488147419103232}, \frac{373859350669998101295}{147573952589676412928} \right], \\ \lambda &= \left[ \frac{330594206547890592307}{4722366482869645213696}, \frac{330594206547890592319}{4722366482869645213696} \right], & b_2 &= \left[ \frac{373859350669998101293}{147573952589676412928}, \frac{23366209416874881331}{9223372036854775808} \right], \\ \lambda &= \left[ \frac{526782486925186168309}{590295810358705651712}, \frac{526782486925186168321}{590295810358705651712} \right], & b_2 &= \left[ \frac{-468997241444025982397}{295147905179352825856}, \frac{-468997241444025982393}{295147905179352825856} \right], \\ \lambda &= \left[ \frac{40465807181488578751}{36893488147419103232}, \frac{323726457451908630039}{295147905179352825856} \right], & b_2 &= \left[ \frac{373859350669998101285}{147573952589676412928}, \frac{373859350669998101303}{147573952589676412928} \right], \\ \lambda &= \left[ \frac{9885258324085039353}{1152921504606846976}, \frac{316328266370721259321}{36893488147419103232} \right], & b_2 &= \left[ \frac{-234498620722012991203}{147573952589676412928}, \frac{-468997241444025982389}{295147905179352825856} \right]. \end{aligned}$$

Une évaluation numérique à 16 chiffres après la virgule (évaluation sous *Maple* avec *Digits := 16*) de ces intervalles donne :

$$\begin{aligned} \lambda &= [-5.725896509837688, -5.725896509837688] & b_2 &= [-1.589024462697880, -1.589024462697880], \\ \lambda &= [-0.1165611381589174, -0.1165611381589174], & b_2 &= [2.533369501252700, 2.533369501252700], \\ \lambda &= [0.07000604627936417, 0.07000604627936417], & b_2 &= [2.533369501252700, 2.533369501252700], \\ \lambda &= [0.8924042449243807, 0.8924042449243807], & b_2 &= [-1.589024462697880, -1.589024462697880], \\ \lambda &= [1.096827901438736, 1.096827901438736], & b_2 &= [2.533369501252700, 2.533369501252700], \\ \lambda &= [8.574094840442733, 8.574094840442733], & b_2 &= [1.589024462697880, -1.589024462697880]. \end{aligned}$$

Nous isolons alors la boîte dont la coordonnée  $\lambda$  est maximale en valeur absolue sachant que la coordonnée  $b_{n-1}$  est réelle maximale :

$$\lambda_a = \left[ \frac{40465807181488578751}{36893488147419103232}, \frac{323726457451908630039}{295147905179352825856} \right], \quad \sigma = \left[ \frac{373859350669998101285}{147573952589676412928}, \frac{373859350669998101303}{147573952589676412928} \right].$$

Une approximation numérique donne alors les valeurs suivantes :

$$\lambda_a = [1.096827901438736, 1.096827901438736], \quad \sigma = [2.533369501252700, 2.533369501252700].$$

Nous obtenons alors la valeur de  $\gamma_{\text{opt}} = \sqrt{1 + \lambda}$  suivante :

$$\gamma_{\text{opt}} = [1.484261245662131, 1.484261245662131].$$

A titre de comparaison, une évaluation des formules symboliques obtenues pour ce même exemple, pour un calcul effectué avec une précision de 20 chiffres après la virgule (évaluation sous *Maple* avec *Digits := 20*), donne :

$$\gamma_{\text{opt}} = 1.4842612456621313162.$$

Toujours à titre de comparaison, la commande *Matlab ncfsyn* appliquée au système *G* donne le résultat suivant :

$$\gamma_{\text{opt}} = 1.484261245662132.$$

Toutes les différentes méthodes donnent donc le même résultat avec une grande précision.

L'avantage d'une résolution via *Matlab* est de ne pas être limité à des systèmes d'ordre faible. En revanche, le calcul numérique n'est pas certifié. Le calcul de  $\gamma_{\text{opt}}$  par une méthode certifiée et l'utilisation de *Maple* fournissent un résultat certifié. En revanche, cette méthode est limitée par l'ordre du système.

### 4.3 Réglage optimal d'une pondération Loop-Shaping

Jusqu'à présent, nous avons proposé une méthode symbolique pour résoudre le problème de commande robuste (PCR) présenté dans la Section 2.3.2. La résolution du PCR est avant tout un moyen permettant de réaliser un asservissement assurant un compromis entre performance et robustesse, ce compromis étant mesurée par  $\gamma_{\text{opt}}$ . Nous pouvons réaliser un asservissement d'un système au moyen de cette approche. Cependant, nous n'avons aucun *moyen de régler* les caractéristiques de l'asservissement obtenu par cette méthode, le PCR ayant une unique solution optimale, autrement dit  $\gamma_{\text{opt}}$  est fixé par les paramètres physiques du système et nous ne pouvons donc pas le régler.

C'est pourquoi le PCR trouve tout son intérêt couplé au concept de Loop-Shaping que nous avons introduit au Chapitre 2. Celui-ci permet d'introduire des *paramètres de réglage* afin de modeler l'asservissement souhaité. Rappelons-en succinctement le principe.

Soit *P* un *système réel*, modèle ne contenant que des paramètres physiques à priori non-modifiables. Celui-ci peut, par exemple, être un système mécanique composé de masses et de ressorts, dont les caractéristiques physiques (masses, raideurs) ne sont pas modifiables. L'approche par Loop-Shaping consiste alors à augmenter ce système réel par une *pondération fréquentielle* *W* afin de créer le *système fictif* *G* :

$$G := P W.$$

Le choix de la pondération *W* se fait au moyen de la procédure décrite dans la Section 2.5. Nous pouvons la voir comme un moyen d'introduire des *paramètres de réglage* permettant de modifier la boucle ouverte de l'asservissement. Le système *G* contient alors des paramètres de réglage. C'est sur ce système que nous travaillons afin de trouver une solution au PCR, c'est-à-dire déterminer  $\gamma_{\text{opt}}$ . L'utilisation de pondérations *W* influe sur la valeur de  $\gamma_{\text{opt}}$  de *G* qui sert d'*indicateur* pour juger si les pondérations choisies sont bonnes ou non. Une idée

serait alors de trouver un moyen de régler ces pondérations de manière automatique. C'est dans ce cadre que l'approche symbolique trouve un autre intérêt.

Dans cette section, nous souhaitons asservir un système à paramètres inconnus par la méthode  $H_\infty$  par Loop-Shaping où la pondération est elle aussi à paramètres inconnus. Dans ce cadre, plus nous introduisons des pondérations avec des dynamiques sophistiquées, plus nous introduisons un grand nombre de paramètres et plus l'ordre du système fictif augmente. Or, dans les sections précédentes, nous avons vu que l'ordre du système et le nombre de paramètres inconnus étaient déterminants quand à la résolution du PCR. C'est pourquoi, dans cette section, nous nous limitons aux choix de *pondérations fréquentielles scalaires de type gain*. Ceci nous permettra de considérer des systèmes réels comportant des dynamiques assez complexes, tout en disposant d'un *moyen de réglage simple* que l'on ajustera suivant la valeur de  $\gamma_{\text{opt}}$ . L'extension de la méthode à d'autres types de pondérations est laissée en perspective.

Passons maintenant à l'explication de la méthode que nous proposons.

### 4.3.1 Une méthode de réglage de pondération scalaire de type gain

Soit  $P$  un système réel dépendant de  $q$  paramètres  $\theta := (\theta_1, \dots, \theta_q)$  non fixés à des valeurs numériques. Supposons que la fonction de transfert de  $P$  soit rationnelle en les paramètres  $\theta$ , c'est-à-dire que  $P(s, \theta) \in \mathbb{Q}(s, \theta)$ . Soit  $w$  un paramètre de réglage qui servira de pondération scalaire. Notons alors  $G = wP$  le système fictif.

Pour toutes valeurs réelles de paramètres  $\theta$  et  $w$ , nous noterons par  $\gamma_{\text{opt}}(w, \theta)$  la solution du PCR associé à  $G$  (rappelons que la définition du PCR est donnée à la Définition 2.7, et que son unique solution est donnée, à paramètres fixés, au Théorème 2.6). Notons que, pour des systèmes SISO  $G$  d'ordre  $n$  avec  $1 \leq n \leq 4$ , nous disposons de formules explicites de  $\gamma_{\text{opt}}$  associé à  $G$ . Voir Section 4.1.1.

Soit  $\bar{\gamma}_{\text{opt}} > 0$ . La valeur  $\bar{\gamma}_{\text{opt}}$  est la valeur de  $\gamma_{\text{opt}}$  que l'on souhaite obtenir après réglage de la pondération  $w$ . Nous adoptons aussi les notations suivantes :

$$\begin{cases} \lambda_m(w, \theta) := \gamma_{\text{opt}}^2(w, \theta) - 1, & (w, \theta) \in \mathbb{R}^2, \\ \bar{\lambda}_m := \bar{\gamma}_{\text{opt}}^2 - 1 \in \mathbb{R}. \end{cases} \quad (4.23)$$

Dans ce contexte, nous définissons le problème suivant.

**Définition 4.2. Problème de réglage de pondération (PRP).**

Trouver une fonction scalaire  $\hat{w}$  de  $\theta \in \mathbb{R}^q$ , c'est-à-dire

$$\begin{aligned} \hat{w} : \mathbb{R}^q &\longrightarrow \mathbb{R} \\ \theta &\longmapsto \hat{w}(\theta), \end{aligned}$$

telle que :

$$\forall \theta \in \mathbb{R}^q, \quad \gamma_{\text{opt}}(\hat{w}(\theta), \theta) = \bar{\gamma}_{\text{opt}}.$$

Afin de résoudre le PRP, supposons que nous disposons d'une paramétrisation rationnelle des solutions  $X \in \mathbb{Q}(b_{n-1}, w, \theta)^{n \times n}$  de l'équation de Riccati algébrique  $\mathcal{R} = 0$  suivant une équation polynomiale  $p = 0$  où  $p \in \mathbb{Q}[b_{n-1}, w, \theta]$ . Une telle paramétrisation peut se trouver pour des systèmes d'ordre faible. Voir la Section 3.6 du Chapitre 3. Disposant de  $X \in$

$\mathbb{Q}(b_{n-1}, w, \theta)^{n \times n}$ , nous pouvons calculer  $Y := \mathcal{Q} X \mathcal{Q} \in \mathbb{Q}(b_{n-1}, w, \theta)^{n \times n}$  (voir Théorème 2.5) ainsi que le numérateur du polynôme caractéristique de  $Y X$ , noté  $h \in \mathbb{Q}[\lambda, b_{n-1}, w, \theta]$ . Nous proposons alors d'étudier la variété algébrique définie par le système polynomial  $\mathcal{T}$  suivant :

$$\mathcal{T} := \begin{cases} p(b_{n-1}, w, \theta) = 0, \\ h(\lambda, b_{n-1}, w, \theta) = 0. \end{cases} \quad (4.24)$$

Le système  $\mathcal{T}$  paramétrise toutes les “valeurs propres”  $\lambda$  de toutes les matrices  $Y X$  possibles, où  $X$  est solution de  $\mathcal{R} = 0$  et  $Y = \mathcal{Q} X \mathcal{Q}$ , y compris les “valeurs propres” de matrices  $X$  ou  $Y$  non définies positives. Comme seul  $\lambda$  est utile ici, puisqu'il permet de calculer directement  $\gamma_{\text{opt}}$ , nous choisissons d'éliminer  $b_{n-1}$  dans  $\mathcal{T}$  en ne gardant que la projection de  $\mathcal{T}$  sur la variable  $\lambda$ . Pour cela, nous proposons de calculer le résultant de  $p$  et  $h$  par rapport à la variable  $b_{n-1}$  (voir Section 3.1.9). Nous obtenons alors le polynôme suivant :

$$L(\lambda, w, \theta) := \text{Res}(p, h, b_{n-1}) = 0. \quad (4.25)$$

Notons que la notion de résultant s'étend aux polynômes à coefficients dans  $\mathbb{Q}(w, \theta)$ , sa définition se basant sur le déterminant d'une matrice de Sylvester.

A paramètres  $w$  et  $\theta$  fixés à des valeurs réelles, une des racines de  $L$  est égale à  $\lambda_m(w, \theta)$  (lié à  $\gamma_{\text{opt}}(w, \theta)$  par (4.23)) et toutes les autres racines de  $L$  correspondent à toutes les autres valeurs propres des matrices  $Y X$ .

Comme nous cherchons une fonction scalaire  $\hat{w}$  des paramètres  $\theta$  telle que

$$\forall \theta \in \mathbb{R}^q, \quad \gamma_{\text{opt}}(\hat{w}(\theta), \theta) = \bar{\gamma}_{\text{opt}} \quad \Leftrightarrow \quad \lambda_m(\hat{w}(\theta), \theta) = \bar{\lambda}_m,$$

nous définissons le polynôme suivant :

$$\bar{L}(w, \theta) := L(\bar{\lambda}_m, w, \theta).$$

$\bar{L}$  est un polynôme multivarié en les variables  $w$  et  $\theta$ . Il définit une courbe dans l'espace  $\mathbb{R}^{q+1}$ . Nous pouvons concevoir la variété  $V(\langle \bar{L} \rangle)$  comme une série de branches. Sur chaque branche, une “valeur propre” de toutes les matrices  $Y X$  possibles est égale à  $\bar{\lambda}_m$ .

Le premier travail est alors d'identifier quelle est la branche qui nous intéresse, c'est-à-dire la branche  $\mathcal{C}$  sur laquelle nous avons :

$$\forall (\theta, w) \in \mathcal{C}, \quad \lambda_m(w, \theta) = \bar{\lambda}_m.$$

Cette dernière tâche peut être réalisée en calculant la variété discriminante de  $\bar{L}$  en  $w$ . Pour cela, définissons le discriminant de  $\bar{L}$  en  $w$  comme suit :

$$\Delta_{\bar{L}}(\theta) = \text{disc}(\bar{L}, w)(\theta).$$

De plus, notons  $\text{LT}(\bar{L}, w)$  le coefficient du terme de plus haut degré en  $w$  de  $\bar{L}$ . Pour un rappel sur les discriminants et les variétés discriminantes, voir la Section 3.1.10. La variété discriminante de  $\bar{L}$  en  $w$  est alors définie par :

$$\mathcal{V}_{\bar{L}} := \{\theta \in \mathbb{C}^q \mid \Delta_{\bar{L}}(\theta) = 0\} \cup \{\theta \in \mathbb{C}^q \mid \text{LT}(\bar{L}, w) = 0\}.$$

La variété  $\mathcal{V}_{\bar{L}}$  décompose l'espace des paramètres défini par  $w$  et  $\theta$  en cellules dans lesquelles le nombre de racines réelles de  $\bar{L}$  en  $w$  est constant. En choisissant un point témoin (déterminé par une valeur rationnelle de  $\theta$ ) dans chacune des cellules, nous pouvons alors isoler les

racines de  $\bar{L}$  en  $w$  de manière certifiée. Nous obtenons alors des couples témoins  $(w, \theta)$  dans chaque cellule régulière permettant d'identifier chaque branche de  $\bar{L}$ . En spécialisant chacune de valeurs témoins dans une expression sous forme close de  $\gamma_{\text{opt}}$ , ou dans  $\mathcal{T}$  puis en isolant ses racines, nous identifions la branche  $\mathcal{L}$  de  $\bar{L}$  telle que :

$$\forall (\theta, w) \in \mathcal{L}, \quad \lambda_m(w, \theta) = \bar{\lambda}_m \quad \Leftrightarrow \quad \gamma_{\text{opt}}(w, \theta) = \bar{\gamma}_{\text{opt}}.$$

**Remarque 4.5.** Notons que cette dernière étape de spécialisation n'est pas certifiée.

Maintenant que nous avons identifié la branche  $\mathcal{L}$  permettant d'obtenir  $\gamma_{\text{opt}}(w, \theta) = \bar{\gamma}_{\text{opt}}$  en tout point  $(w, \theta)$  de la branche, il reste à définir une méthode pour définir  $\hat{w}$  permettant de rester sur cette branche  $\mathcal{L}$  quelque soit  $\theta$ . Disposant de la variété discriminante de  $\bar{L}$  et de la décomposition en cellules associée, nous distinguons deux cas suivant le degré du polynôme  $\bar{L}$ .

- Dans le premier cas, calculer les racines du polynôme  $\bar{L}$  en  $w$  est équivalent à trouver les racines d'un polynôme de degré inférieur ou égal à 4. Dans ce cas, nous pouvons expliciter ses racines par radicaux en utilisant l'Annexe D. Nous disposons alors d'expressions par radicaux de  $w$  solution de  $\bar{L} = 0$  dépendant de  $\theta$ , notées  $\hat{w}_i$ , et l'une d'elles permet d'atteindre  $\bar{\gamma}_{\text{opt}}$ , si celui-ci est atteignable. Considérons alors l'union des variétés singulières où les expressions par radicaux  $\hat{w}_i$  ne sont pas définies et la variété discriminante de  $\bar{L}$  en  $w$ . Ces variétés décomposent l'espace des paramètres en cellules régulières  $\mathcal{C}_i$  dans lesquelles le nombre de racines réelles de  $\bar{L}$  en  $w$  est constant, et où les expressions par radicaux  $\hat{w}_i$  sont continues. L'identification de l'expression par radicaux telle que

$$\forall \theta \in \mathcal{C}_i, \quad \gamma_{\text{opt}}(\hat{w}_i(\theta), \theta) = \bar{\gamma}_{\text{opt}},$$

peut alors se faire par évaluation numérique de  $\gamma_{\text{opt}}(\hat{w}_i(\theta^c), \theta^c)$  en un point témoin  $\theta^c$  de chaque cellule  $\mathcal{C}_i$ .

- Dans le cas contraire, nous ne pouvons trouver de racines par radicaux de  $L$  en  $w$ . Nous proposons alors d'approximer la branche permettant d'atteindre  $\bar{\gamma}_{\text{opt}}$  localement, c'est-à-dire autour d'un point de fonctionnement par exemple, par un développement en série de Taylor ou de Puiseux [BPR03, Chap. 6.2]. Ce type de développement fournit une fonction scalaire :

$$\begin{aligned} \hat{w} : \mathbb{R}^q &\longrightarrow \mathbb{R} \\ \theta &\longmapsto \hat{w}(\theta). \end{aligned}$$

En choisissant comme pondération  $\hat{w}(\theta)$ , nous assurons alors que  $\gamma_{\text{opt}}$  est proche de  $\bar{\gamma}_{\text{opt}}$  au voisinage du point de fonctionnement choisi.

Ces deux cas sont illustrés sur deux exemples dans les sous-sections qui suivent. Afin de mieux représenter les variétés entrant en jeu et de simplifier les calculs, notons que les exemples qui suivent portent sur des systèmes avec au plus 2 paramètres. Ceci nous permettra de réaliser des tracés dans l'espace, ce qui devient difficilement réalisable avec un plus grand nombre de paramètres (à moins de réaliser des coupes).

**Remarque 4.6.** Si aucune branche de  $\bar{L}$  ne correspond à  $\lambda_m$ , alors la valeur  $\bar{\gamma}_{\text{opt}}$  n'est pas atteignable par  $\gamma_{\text{opt}}$ . Notamment, les valeurs limites atteignables par  $\lambda_m$  sont contenues dans la variété discriminante de  $L$  en  $w$ .

**Remarque 4.7.** Nous avons considéré ici le système  $\mathcal{T}$  défini par  $p$  et le polynôme caractéristique de  $YX$ . Nous aurions pu appliquer le même genre de démarche au système  $\mathcal{T}_*$ ,

défini par  $p$  et le numérateur du polynôme caractéristique de  $\mathcal{Q}X$  noté  $h_*$  :

$$\mathcal{T}_* := \begin{cases} p(b_{n-1}, w, \theta) = 0, \\ h_*(\lambda, b_{n-1}, w, \theta) = 0. \end{cases}$$

La démarche resterait la même à la différence près que nous devons faire attention à la valeur absolue de  $\lambda$ . En effet, si nous considérons  $\mathcal{T}_*$ , alors  $\gamma_{\text{opt}}$  est défini à partir de  $\lambda_*(\mathcal{Q}X)$ , c'est-à-dire la valeur absolue maximale des valeurs propres réelles de  $\mathcal{Q}X$ . Nous devons alors considérer les deux systèmes suivants :

$$\mathcal{T}_*^+ := \begin{cases} p(b_{n-1}, w, \theta) = 0, \\ h_*(\lambda, b_{n-1}, w, \theta) = 0, \end{cases} \quad \mathcal{T}_*^- := \begin{cases} p(b_{n-1}, w, \theta) = 0, \\ h_*(-\lambda, b_{n-1}, w, \theta) = 0. \end{cases}$$

La projection de chacun de ces systèmes sur  $\lambda$  donne alors :

$$L^+(\lambda, w, \theta) := \text{Res}(p, h_*(\lambda), b_{n-1}) = 0, \quad L^-(\lambda, w, \theta) := \text{Res}(p, h_*(-\lambda), b_{n-1}) = 0.$$

Chacun de ces polynômes constitue une courbe dans l'espace des paramètres. Il faut alors étudier l'union de ces deux variétés. Notamment, la décomposition en cellules devra prendre en compte les intersections de ces deux variétés, une intersection correspondant au lieu où une racine de  $\mathcal{Q}X$  est égale à l'opposée d'une autre racine de  $\mathcal{Q}X$ .

### 4.3.2 Application au système double masse-ressort

Reprenons l'exemple du système masse-ressort d'ordre 4 développé dans la Section 4.1.3 et représenté par la Figure 4.13.

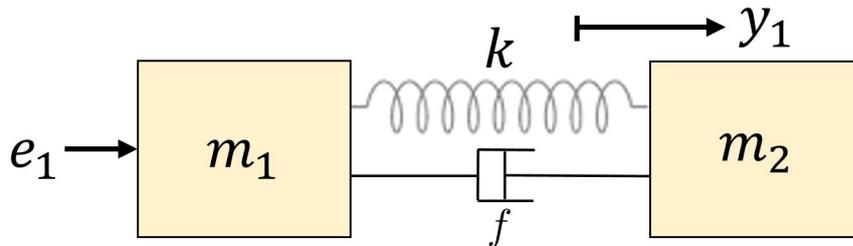


FIGURE 4.13 – Système deux masses ressort amorti

Considérons dans un premier temps le cas sans amortissement, c'est-à-dire  $f = 0$ , et montrons que nous pouvons trouver une expression explicite d'une pondération Loop-Shaping scalaire permettant d'atteindre un  $\gamma_{\text{opt}}$  donné.

#### Étude du système sans amortissement : réglage de pondération global

Dans la Section 4.1.3, nous avons conclu que, pour le système sans amortissement, l'expression sous forme close de  $\gamma_{\text{opt}}$  était très facile à obtenir et pouvait se calculer à la main, de nombreuses simplifications et symétries apparaissant dans les polynômes étudiés. Dans un premier temps, considérons alors  $f = 0$ .

Soit une pondération scalaire de type gain  $w$ . Nous définissons le *système réel*  $P$  par

$$P := \frac{c'_0}{s^2 (s^2 + a_2)},$$

et le *système fictif*  $G$  par

$$G := w P = \frac{c_0}{s^2 (s^2 + a_2)}, \quad (4.26)$$

où :

$$c_0 := c'_0 w.$$

De plus, notons par  $\theta$  les paramètres de  $G$ , c'est-à-dire :

$$\theta := (c_0, a_2).$$

Reprenons les notations de la Section 4.1.3, c'est-à-dire :

$$\begin{cases} \sigma := \sqrt{2} \sqrt{(\sqrt{2} c_0 - a_2) + \sqrt{(\sqrt{2} c_0 - a_2)^2 + 2 c_0}}, \\ t_+ := \frac{1}{8 c_0} \left( \sigma^4 + 8 c_0 + \sqrt{(\sigma^4 + 8 c_0)^2 + 32 c_0 \sigma^4} \right), \\ \lambda_a := \frac{1}{2} \left( t_+ + \sqrt{t_+^2 + 4} \right). \end{cases} \quad (4.27)$$

La Section 4.1.3 montre que la formule explicite de  $\gamma_{\text{opt}}$  est donnée par :

$$\forall (c_0, a_2) \in ]0, +\infty[^2, \quad \gamma_{\text{opt}}(c_0, a_2) := \sqrt{1 + \lambda_a^2}. \quad (4.28)$$

Soit  $\bar{\gamma}_{\text{opt}} > 0$ . L'objectif fixé est le suivant.

**Définition 4.3. Problème du réglage global de la pondération**

Trouver une fonction scalaire  $\hat{c}_0$  de  $a_2 \in \mathbb{R}_+$ , c'est-à-dire

$$\begin{aligned} \hat{c}_0 : \mathbb{R}_+ &\longrightarrow \mathbb{R}_+ \\ a_2 &\longmapsto \hat{c}_0(a_2), \end{aligned}$$

telle que :

$$\forall a_2 \in \mathbb{R}_+, \quad \gamma_{\text{opt}}(\hat{c}_0(a_2), a_2) = \bar{\gamma}_{\text{opt}}.$$

Remarquons alors dans un premier temps la propriété suivante.

**Proposition 4.2.** Soit  $(c_0, a_2) \in ]0, +\infty[^2$ . Alors, nous avons

$$\gamma_{\text{opt}}(c_0, a_2) > \sqrt{2},$$

et :

$$\forall \bar{\gamma}_{\text{opt}} > \sqrt{2}, \quad \gamma_{\text{opt}}(c_0, a_2) = \bar{\gamma}_{\text{opt}} \Leftrightarrow t_+(c_0, a_2) = \frac{\bar{\gamma}_{\text{opt}}^2 - 2}{\sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}}.$$

*Démonstration.* La Section 4.1.3 montre la propriété suivante :

$$\forall (c_0, a_2) \in ]0, +\infty[^2, \quad t_+(c_0, a_2) > 0.$$

Ainsi, en utilisant (4.27), nous en déduisons

$$\forall (c_0, a_2) \in ]0, +\infty[^2, \quad \lambda_a(c_0, a_2) > 1,$$

et :

$$\forall (c_0, a_2) \in ]0, +\infty[^2, \quad \gamma_{\text{opt}}(c_0, a_2) > \sqrt{2}.$$

Soient  $\bar{\gamma}_{\text{opt}} > \sqrt{2}$  et  $(c_0, a_2) \in ]0, +\infty[^2$  tels que :

$$\gamma_{\text{opt}}(c_0, a_2) = \bar{\gamma}_{\text{opt}}. \quad (4.29)$$

En reprenant les notations (4.27), l'équation (4.29) est équivalente à :

$$\lambda_a(c_0, a_2) = \pm \sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}. \quad (4.30)$$

En mettant au carré chaque côté de l'équation (4.30), nous obtenons :

$$\lambda_a(c_0, a_2)^2 = \bar{\gamma}_{\text{opt}}^2 - 1. \quad (4.31)$$

Comme  $\bar{\gamma}_{\text{opt}} > \sqrt{2}$ , alors (4.31) est équivalente à :

$$\lambda_a(c_0, a_2) = \sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}. \quad (4.32)$$

L'équation (4.32) est alors équivalente à

$$\frac{1}{2} \left( t_+(c_0, a_2) + \sqrt{t_+(c_0, a_2)^2 + 4} \right) = \sqrt{\bar{\gamma}_{\text{opt}}^2 - 1},$$

c'est-à-dire :

$$\sqrt{t_+(c_0, a_2)^2 + 4} = 2 \sqrt{\bar{\gamma}_{\text{opt}}^2 - 1} - t_+(c_0, a_2). \quad (4.33)$$

En mettant (4.33) au carré, nous obtenons alors

$$t_+(c_0, a_2) = \frac{\bar{\gamma}_{\text{opt}}^2 - 2}{\sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}}, \quad (4.34)$$

ce qui conclut la preuve. □

La Proposition 4.2 montre que chercher une fonction scalaire  $\hat{c}_0$  de  $a_2 \in \mathbb{R}_+$  telle que

$$\forall \bar{\gamma}_{\text{opt}} > \sqrt{2}, \quad \forall a_2 \in \mathbb{R}_+, \quad \gamma_{\text{opt}}(\hat{w}(a_2), a_2) = \bar{\gamma}_{\text{opt}},$$

est équivalent à chercher une fonction scalaire  $\hat{c}_0$  de  $a_2 \in \mathbb{R}_+$  telle que :

$$\forall \bar{t}_+ > 0, \quad \forall a_2 \in \mathbb{R}_+, \quad t_+(\hat{w}(a_2), a_2) = \bar{t}_+.$$

Dans la suite, nous notons alors :

$$\bar{t}_+ := \frac{\bar{\gamma}_{\text{opt}}^2 - 2}{\sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}}.$$

D'autre part, la Section 4.1.3 (voir (4.11) et (4.14)) montre que  $t_+$  est solution du système polynomial suivant :

$$\mathcal{T}_\star := \begin{cases} p_\tau(\tau) = c_0 \tau^4 + \sqrt{8c_0} a_2 \tau^3 + 2(a_2^2 - 3c_0) \tau^2 - \sqrt{8c_0} a_2 \tau + c_0 = 0, \\ h_s(t, \tau) = 4c_0 t^2 - (\tau^2 + 8c_0)t - 2\tau^2. \end{cases} \quad (4.35)$$

Les racines en  $t$  du système  $\mathcal{T}_\star$  défini par (4.35) permettent de calculer les valeurs propres d'une partie des matrices  $\mathcal{L}X$  possibles, où  $X$  est solution de  $\mathcal{R} = 0$ , y compris les valeurs propres de matrices  $X$  non définies positives. Comme seul  $t$  est utile ici, puisqu'il permet de calculer directement  $\gamma_{\text{opt}}$ , nous choisissons d'éliminer  $\tau$  dans  $\mathcal{T}_\star$  en ne gardant que la projection de  $\mathcal{T}_\star$  sur la variable  $t$ . Pour cela, nous proposons de calculer le résultant de  $p_\tau$  et  $h_s$  par rapport à la variable  $\tau$  (voir Section 3.1.9). Nous obtenons alors le polynôme suivant :

$$L_t := (t^4 - 16t^3 + 8t^2 + 64t + 16)^2 c_0^2 - 8t(t-2)(t+2)(t^2+4)^2 a_2^2 c_0 + 16t^2(t-2)^2(t+2)^2 a_2^4. \quad (4.36)$$

Notons que la notion de résultant s'étend aux polynômes à coefficients dans  $\mathbb{Q}(c_0, a_2)$ , sa définition se basant sur le déterminant d'une matrice de Sylvester. À paramètres  $c_0$  et  $a_2$  fixés à des valeurs réelles, une des racines de  $L_t$  est égale à  $t_+(c_0, a_2)$  (liée à  $\gamma_{\text{opt}}$  par (4.27)) et toutes les autres racines de  $L_t$  correspondent à d'autres valeurs propres des matrices  $\mathcal{L}X$ .

Notons  $\Delta$  le discriminant de  $L_t$  en  $c_0$ , c'est-à-dire

$$\Delta := 2048 t^3 a_2^4 (t^2 - 4t - 4)^2 (t - 2)^3 (t + 2)^3,$$

et  $\text{LT}(L_t, c_0)$  le terme de tête de  $L_t$  en  $c_0$  :

$$\text{LT}(L_t, c_0) = (t^4 - 16t^3 + 8t^2 + 64t + 16)^2.$$

Le lieu d'annulation de ces deux polynômes constitue la variété discriminante de  $L_t$  en  $c_0$ . Celle-ci divise alors l'espace des paramètres en cellules dans lesquelles le nombre de racines réelles de  $L_t$  en  $c_0$  est constant. Notamment, nous avons

$$\begin{aligned} \{(a_2, t) \in \mathbb{R}_+^2 \mid \Delta = 0\} &= \{(a_2, t) \in \mathbb{R}_+^2 \mid a_2 = 0\} \cup \{(a_2, t) \in \mathbb{R}_+^2 \mid t = 2\} \\ &\cup \{(a_2, t) \in \mathbb{R}_+^2 \mid t = 2 + 2\sqrt{2} \approx 4.82\}, \\ &\cup \{(a_2, t) \in \mathbb{R}_+^2 \mid t = 0\}, \end{aligned}$$

et :

$$\begin{aligned} \{(a_2, t) \in \mathbb{R}_+^2 \mid \text{LT}(L_t, c_0) = 0\} &= \{(a_2, t) \in \mathbb{R}_+^2 \mid t = 4 + 2\sqrt{6} - 2\sqrt{3} - 2\sqrt{2} \approx 2.60\} \\ &\cup \{(a_2, t) \in \mathbb{R}_+^2 \mid t = 4 + 2\sqrt{6} + 2\sqrt{3} + 2\sqrt{2} \approx 15.2\}. \end{aligned}$$

Toutes ces valeurs délimitent l'espace des paramètres  $\mathbb{R}_+^2$  en cellules  $\mathcal{C}_i$  dans lesquelles le nombre de racines réelles de  $L_t$  en  $c_0$  est constant. Dans la suite, nous notons :

$$\begin{cases} t_f := 2\sqrt{2} + 2 \approx 4.82, \\ t_{\min} := 4 + 2\sqrt{6} - 2\sqrt{3} - 2\sqrt{2} \approx 2.60, \\ t_{\sup} := 4 + 2\sqrt{6} + 2\sqrt{3} + 2\sqrt{2} \approx 15.2. \end{cases}$$

$L_t(t, c_0, a_2)$  est un polynôme de degré 2 en  $c_0$ . Ses racines en  $c_0$  peuvent alors s'expliciter par radicaux (voir Annexe D). Celles-ci s'écrivent

$$\forall \varepsilon \in \{-1, 1\}, \quad C_0(\varepsilon)(t, a_2) := \frac{8t(t-2)(t+2)(t^2+4)^2 a_2^2 + \varepsilon \sqrt{\Delta}}{2(t^4 - 16t^3 + 8t^2 + 64t + 16)^2},$$

c'est-à-dire ( $a_2 > 0$ ), pour  $\varepsilon \in \{-1, 1\}$  :

$$C_0(\varepsilon)(t, a_2) = \frac{4t(t-2)(t+2)(t^2+4)^2 a_2^2 + \varepsilon 16\sqrt{2} a_2^2 \sqrt{t^3(t^2-4t-4)^2(t-2)^3(t+2)^3}}{(t^4 - 16t^3 + 8t^2 + 64t + 16)^2}. \quad (4.37)$$

**Remarque 4.8.** Dans une cellule  $\mathcal{C}_i$ , les deux expressions  $C_0(1)$  et  $C_0(-1)$  sont continues et ne sont jamais égales. Cette propriété découle directement de la définition de la variété discriminante  $L_t$  en  $c_0$ .

Considérons alors une cellule  $\mathcal{C}_i$  de la décomposition et un point de  $\mathcal{C}_i$ , noté  $(t^c, a_2^c)$ . Soit  $\varepsilon \in \{-1, 1\}$ . Considérons alors la procédure suivante.

1. Évaluer  $C_0(\varepsilon)(t, a_2)$  défini par (4.37) en  $(t^c, a_2^c)$ .
  - Si  $C_0(\varepsilon)(t^c, a_2^c) \in \mathbb{R}$ , alors le point  $(t^c, C_0(\varepsilon)(t^c, a_2^c), a_2^c)$  est sur la variété  $L_t = 0$ . Par conséquent, pour tout  $(t, a_2) \in \mathcal{C}_i$ , alors  $(t, C_0(\varepsilon)(t, a_2), a_2)$  est sur la variété  $L_t = 0$  car  $C_0(\varepsilon)$  est continue sur  $\mathcal{C}_i$ .
  - Dans le cas contraire,  $L_t$  n'a pas de racine réelle en  $c_0$  sur  $\mathcal{C}_i$ . Arrêter la procédure.
2. Évaluer  $t_+(C_0(\varepsilon), a_2^c)$  (voir (4.27)).  
 Si  $t_+(C_0(\varepsilon), a_2^c) = t^c$ , alors la fonction scalaire  $\hat{c}_0 = C_0(\varepsilon)$  permet d'atteindre  $t_+ = \bar{t}_+$  pour tout  $(a_2, \bar{t}_+) \in \mathcal{C}_i$  car  $C_0(\varepsilon)$  est continue sur  $\mathcal{C}_i$ .

Nous pouvons alors appliquer cette procédure à toutes les cellules  $\mathcal{C}_i$  et obtenir le graphe donné dans la Figure 4.14. Pour  $\bar{t}_+ \in [t_f, t_{\max}]$  (cellule verte), le choix de la fonction scalaire  $\hat{c}_0 = C_0(1)(\bar{t}_+, a_2)$  permet d'atteindre la valeur de  $\bar{t}_+$  désirée. Pour  $\bar{t}_+ \in [t_{\min}, t_f]$  (cellule bleue), le choix de la fonction scalaire  $\hat{c}_0 = C_0(-1)(\bar{t}_+, a_2)$  permet d'atteindre la valeur de  $\bar{t}_+$  désirée.

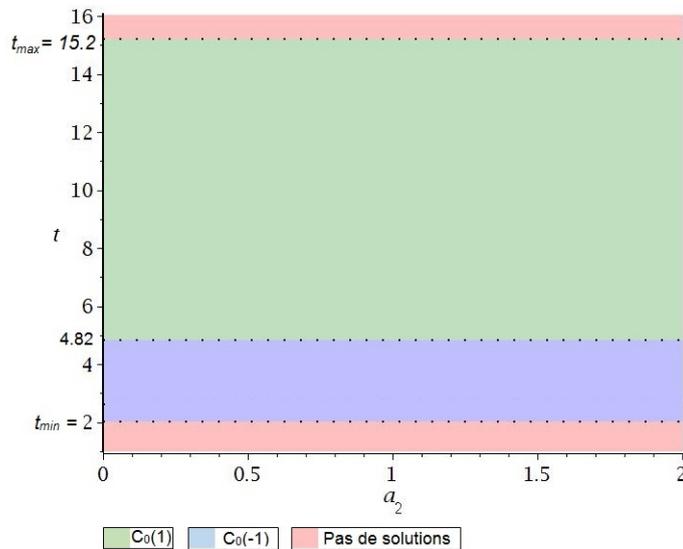


FIGURE 4.14 – Décomposition de l'espace des paramètres à l'aide de la variété discriminante de  $L_t$

**Remarque 4.9.** Cette étude montre que  $\gamma_{\text{opt}}$  est compris entre les deux valeurs suivantes :

$$\begin{aligned} \gamma_{\min} &:= \sqrt{1 + \left( \frac{1}{2} \left( t_{\min} + \sqrt{t_{\min}^2 + 4} \right) \right)^2} = \sqrt{2\sqrt{2} + 4} \approx 2.61, \\ \gamma_{\max} &:= \sqrt{1 + \left( \frac{1}{2} \left( t_{\max} + \sqrt{t_{\max}^2 + 4} \right) \right)^2} \\ &= \sqrt{1 + \left( 2 + \sqrt{6} + \sqrt{3} + \sqrt{2} + \frac{1}{2} \sqrt{64 + (8\sqrt{3} + 8\sqrt{2} + 16)\sqrt{6} + (8\sqrt{2} + 16)\sqrt{3} + 16\sqrt{2}} \right)^2} \\ &\approx 15.29. \end{aligned} \tag{4.38}$$

Récapitulons alors les résultats dans la proposition suivante.

**Proposition 4.3.** Soient  $\bar{\gamma}_{\text{opt}} \in [\gamma_{\min}, \gamma_{\max}]$ ,  $\gamma_f := \sqrt{1 + \left( 1 + \sqrt{2} + \sqrt{4 + 2\sqrt{2}} \right)^2} \approx 5.126$  et  $C_0$  défini en (4.37).

- Si  $\gamma_{\min} < \bar{\gamma}_{\text{opt}} \leq \gamma_f$ , alors en choisissant la pondération

$$w := \frac{1}{c'_0} C_0(-1),$$

la valeur de  $\gamma_{\text{opt}}$  associée au système  $G$  tend vers la valeur souhaitée, c'est-à-dire, en notant  $\bar{t}_+ := \frac{\bar{\gamma}_{\text{opt}}^2 - 2}{\sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}}$ , nous avons :

$$\gamma_{\text{opt}}(C_0(-1)(\bar{t}_+, a_2), a_2) = \bar{\gamma}_{\text{opt}}.$$

- Si  $\gamma_f \leq \bar{\gamma}_{\text{opt}} \leq \gamma_{\max}$ , alors en choisissant la pondération

$$w := \frac{1}{c'_0} C_0(1),$$

la valeur de  $\gamma_{\text{opt}}$  associée au système  $G$  tend vers la valeur souhaitée, c'est-à-dire, en notant  $\bar{t}_+ := \frac{\bar{\gamma}_{\text{opt}}^2 - 2}{\sqrt{\bar{\gamma}_{\text{opt}}^2 - 1}}$ , nous avons :

$$\gamma_{\text{opt}}(C_0(1)(\bar{t}_+, a_2), a_2) = \bar{\gamma}_{\text{opt}}.$$

**Exemple 4.2.** Par exemple, la Proposition 4.3 permet d'affirmer que  $\gamma_{\text{opt}} = 3$  est atteignable en choisissant :

$$\begin{aligned} w &= \frac{4a_2^2}{197261828370049c'_0} \left( 6246072 + 3625216\sqrt{2} - \sqrt{90755864194240 + 67218229682478\sqrt{2}} \right)^2, \\ &\approx 0.103 \left( \frac{a_2^2}{c'_0} \right). \end{aligned}$$

Examinons maintenant le cas avec amortissement.

### Étude du système avec amortissement : réglage local de pondération

Nous proposons maintenant d'ajouter un paramètre indéterminé. Reprenons toujours l'exemple du système masse-ressort d'ordre 4 développé dans la Section 4.1.3 mais cette fois-ci pour un amortissement non nul.

Soit une pondération scalaire de type gain  $w$ . Nous définissons le *système réel*  $P$  par

$$P := \frac{1}{m s^2} \frac{a_3 s + a_2}{(s^2 + a_3 s + a_2)},$$

et le *système fictif*  $G$  par :

$$\frac{y_1}{\tilde{e}_1} := G = w P = \frac{w}{m s^2} \frac{a_3 s + a_2}{(s^2 + a_3 s + a_2)}. \quad (4.39)$$

De plus, notons par  $\theta$  les paramètres de  $P$ , c'est-à-dire :

$$\theta := (m, a_3, a_2).$$

En Section 4.1.3, nous avons déterminé une formule explicite de  $\gamma_{\text{opt}}$  pour le système  $P$ . Passer du système réel au système fictif se fait en multipliant le système réel par une pondération scalaire. Ainsi, seul le gain du système est modifié entre ces deux systèmes. Disposant d'une formule explicite de  $\gamma_{\text{opt}}$  de  $P$  dépendant des paramètres  $\theta := (m, a_3, a_2)$ , il suffit alors de remplacer  $m$  par  $m/w$  dans l'expression de  $\gamma_{\text{opt}}$  pour obtenir une expression explicite de  $\gamma_{\text{opt}}$  pour le système fictif  $G$  dépendant de  $(w, \theta)$ . L'expression explicite de  $\gamma_{\text{opt}}$  de  $G$  ne dépend alors que de deux paramètres  $G_r$  et  $\rho$  définis par :

$$\begin{cases} G_r := \frac{w}{m a_2}, \\ \rho := \frac{a_3}{\sqrt{a_2}}. \end{cases} \quad (4.40)$$

Afin de ne pas alourdir les notations, nous noterons  $\gamma_{\text{opt}}(G_r, \rho)$  cette nouvelle expression explicite de  $\gamma_{\text{opt}}$ , bien que celle-ci soit différente de celle exprimée avec les paramètres  $(w, \theta)$ .

En pratique, les ingénieurs automaticiens veulent concevoir un asservissement permettant d'assurer un bon compromis performance/robustesse. Un tel compromis est en général obtenu pour des valeurs de  $\gamma_{\text{opt}}$  entre 2 et 3. Dans cet exemple, nous choisirons arbitrairement qu'une bonne valeur de  $\gamma_{\text{opt}}$  est :

$$\bar{\gamma}_{\text{opt}} := 3.$$

Une telle valeur de  $\gamma_{\text{opt}}$  assure en effet des marges de stabilité relativement confortables :  $\Delta_G \geq 6$  dB,  $\Delta_\Phi \geq 39^\circ$  (voir (2.10) de la Remarque 2.6).

L'objectif que l'on se fixe est le suivant.

#### Définition 4.4. Problème du réglage local de la pondération

Trouver une fonction scalaire  $\widehat{G}_r$  de  $\rho \in \mathbb{R}$ , c'est-à-dire

$$\begin{aligned} \widehat{w} : \mathbb{R}_+ &\longrightarrow \mathbb{R}_+ \\ \rho &\longmapsto \widehat{G}_r(\rho), \end{aligned}$$

telle que :

$$\forall \rho \in \mathbb{R}, \quad \gamma_{\text{opt}}(\widehat{G}_r(\rho), \rho) = \bar{\gamma}_{\text{opt}}.$$

Notons que, si l'on trouve une expression  $\widehat{G}_r$  de  $G_r$  permettant d'atteindre  $\bar{\gamma}_{\text{opt}}$  quel que soit  $\rho$ , alors le choix de la fonction scalaire  $\widehat{w} := m a_2 \widehat{G}_r$  comme pondération permet d'obtenir :

$$\forall(\theta) \in \mathbb{R}^3, \quad \gamma_{\text{opt}}(\widehat{w}(\theta), \theta) = \bar{\gamma}_{\text{opt}}.$$

Dans la suite, nous raisonnerons alors sur les paramètres réduits  $G_r$  et  $\rho$ , ainsi que sur l'expression sous forme close de  $\gamma_{\text{opt}}(G_r, \rho)$  dépendant de ces deux paramètres.

La Section 4.1.3 définit un système de deux polynômes  $\tilde{p}$  et  $h$  dont un des zéros est associé à la solution du PCR appliqué au système réel  $P$ . En remplaçant  $m$  par  $m/w$  dans ces polynômes, nous obtenons les deux polynômes suivants :

$$\begin{cases} \tilde{p}_w(b_3, b_0, w, \theta) = 0, \\ h_w(\lambda, b_3, w, \theta) = 0, \end{cases}$$

Ces deux polynômes  $\tilde{p}_w$  et  $h_w$  définissent un système dont un des zéros est la solution du PCR appliqué au système fictif  $G$ . De plus, dans cet exemple, rappelons que nous avons :

$$b_0 = \frac{a_2 w}{m}. \quad (4.41)$$

Ainsi, en spécialisant  $b_0$  par cette dernière expression (4.41) dans  $\tilde{p}_w$ , le polynôme  $\tilde{p}_w$  reste à coefficients rationnels en les paramètres et nous pouvons utiliser toutes les techniques d'élimination vues dans les chapitres précédents. Le système polynomial obtenu après spécialisation de  $b_0$  (4.41) est noté comme suit :

$$\tilde{\mathcal{T}} := \begin{cases} \tilde{p}_w(b_3, w, \theta) = 0, \\ h_w(\lambda, b_3, w, \theta) = 0. \end{cases} \quad (4.42)$$

Par ailleurs, notons  $\lambda_m$  l'expression par radicaux de la racine réelle maximale de  $h$  en  $\lambda$ .

Les racines en  $\lambda$  du système  $\tilde{\mathcal{T}}$  défini par (4.42) permettent de calculer les "valeurs propres" d'une partie des matrices  $Y X$  possibles, où  $X$  est solution de  $\mathcal{R} = 0$ , y compris les valeurs propres de matrices  $X$  et  $Y$  non définies positives. Comme seul  $\lambda$  est utile ici, puisqu'il permet de calculer directement  $\lambda_m$ , nous choisissons d'éliminer  $b_3$  dans  $\tilde{\mathcal{T}}$  en ne gardant que la projection de  $\tilde{\mathcal{T}}$  sur la variable  $\lambda$ . Pour cela, nous proposons de calculer le résultant de  $\tilde{p}_w$  et  $h_w$  par rapport à la variable  $b_3$  (voir Section 3.1.9). Nous obtenons alors le polynôme suivant :

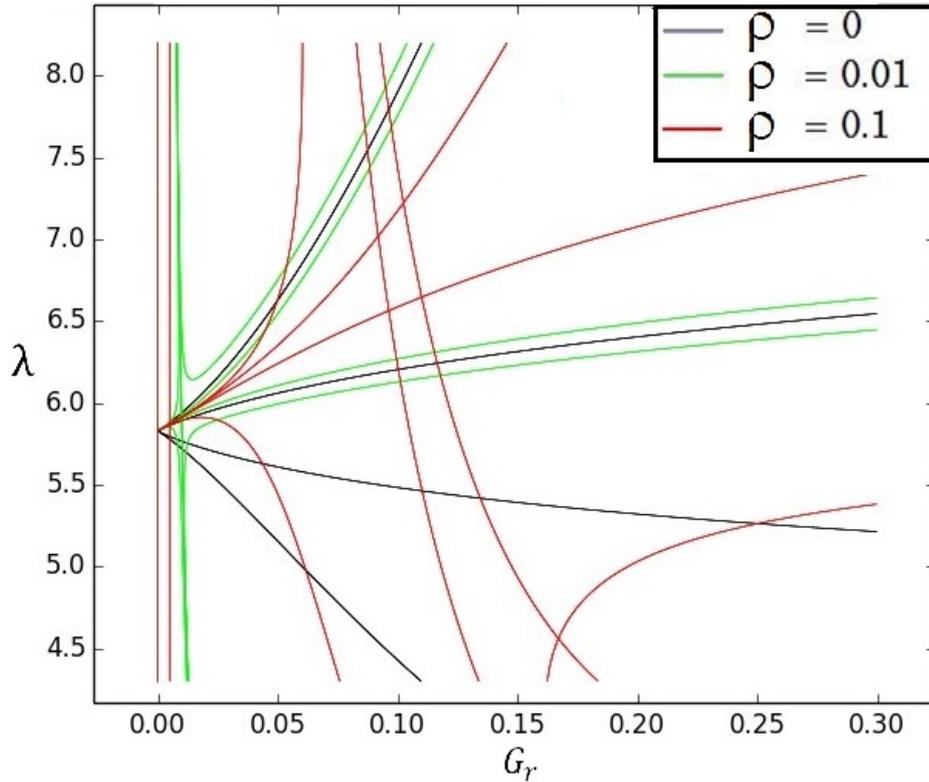
$$\mathcal{L}(\lambda, w, \theta) = \text{Res}(\tilde{p}, h, b_3) = 0. \quad (4.43)$$

A paramètres  $(w, \theta)$  fixés à des valeurs réelles, une des racines de  $\mathcal{L}$  est égale à  $\lambda_m$  et toutes les autres racines de  $\mathcal{L}$  correspondent à d'autres valeurs propres des matrices  $Y X$ .

Le polynôme  $\tilde{p}_w$  est de degré 8 en  $b_3$  et  $h_w$  est de degré 4 en  $\lambda$ . Ainsi,  $\mathcal{L}$  est de degré 32 en  $\lambda$ . Ce dernier est bien trop grand pour être affiché ici. Cependant, il reste exploitable par des logiciels de calcul formel. Par ailleurs, nous remarquons que  $\mathcal{L}(\lambda, \theta)$  ne dépend que de deux paramètres  $G_r$  et  $\rho$  définis en (4.40), tout comme l'expression par radicaux de  $\gamma_{\text{opt}}$ . En effectuant les changements de variable permettant de remplacer les paramètres  $(w, \theta)$  par  $G_r$  et  $\rho$ , nous obtenons le polynôme projeté suivant :

$$L(\lambda, G_r, \rho) = 0.$$

Un tracé de  $L$  pour  $\rho \in \{0, 1/100, 1/10\}$  est donné dans la Figure 4.15.


 FIGURE 4.15 –  $L(\lambda, G_r, \rho) = 0$  pour  $\rho \in \{0, 1/100, 1/10\}$ 

**Remarque 4.10.** Notons qu'un tel tracé est extrêmement difficile à obtenir, et que même les traceurs de *Maple* et des autres logiciels dédiés à l'affichage de courbes ne permettent pas d'obtenir des courbes précises. Ce tracé a été obtenu grâce à une collaboration avec Fabrice Rouillier. Celle-ci nous a permis de tracer  $L$  avec la précision souhaitée. Une courbe animée présentant  $L$  avec  $\rho$  variant de 0 à  $3/4$  a aussi été obtenue, voir [https://who.rocq.inria.fr/Fabrice.Rouillier/DARES/ifac\\_2017.html](https://who.rocq.inria.fr/Fabrice.Rouillier/DARES/ifac_2017.html).

Comme l'on est uniquement intéressé par les valeurs de  $(G_r, \rho)$  donnant  $\gamma_{\text{opt}}(G_r, \rho) = \bar{\gamma}_{\text{opt}}$ , c'est-à-dire  $\lambda_m(G_r, \rho) = \bar{\lambda}_m$  où

$$\bar{\lambda}_m := \bar{\gamma}_{\text{opt}}^2 - 1 = 8,$$

nous définissons :

$$\bar{L}(G_r, \rho) := L(\bar{\lambda}_m, G_r, \rho).$$

La Figure 4.16 représente  $\bar{L} = 0$ .

La variété  $\bar{L} = 0$  consiste en une série de branches. Sur chaque branche, une des "valeurs propres" des matrices  $Y X$  possibles est égale à  $\bar{\lambda}_m$ . Nous voulons identifier sur quelle branche  $\lambda_m = \bar{\lambda}_m$ .

Pour cela, calculons la variété discriminante de  $\bar{L}$  en  $G_r$ , notée  $\Delta_{\bar{L}}$ . Celle-ci décompose l'espace des paramètres défini par  $\rho$  en cellules  $\mathcal{C}_i$  dans lesquelles le nombre de racines réelles de  $\bar{L}$  en  $G_r$  est constant (chaque cellule  $\mathcal{C}_i$  est donc un intervalle délimitant des valeurs de  $\rho$ ). Nous pouvons alors choisir un point témoin  $\rho^c$  dans  $\mathcal{C}_i$ . En évaluant  $\bar{L}$  en  $\rho^c$ , puis en

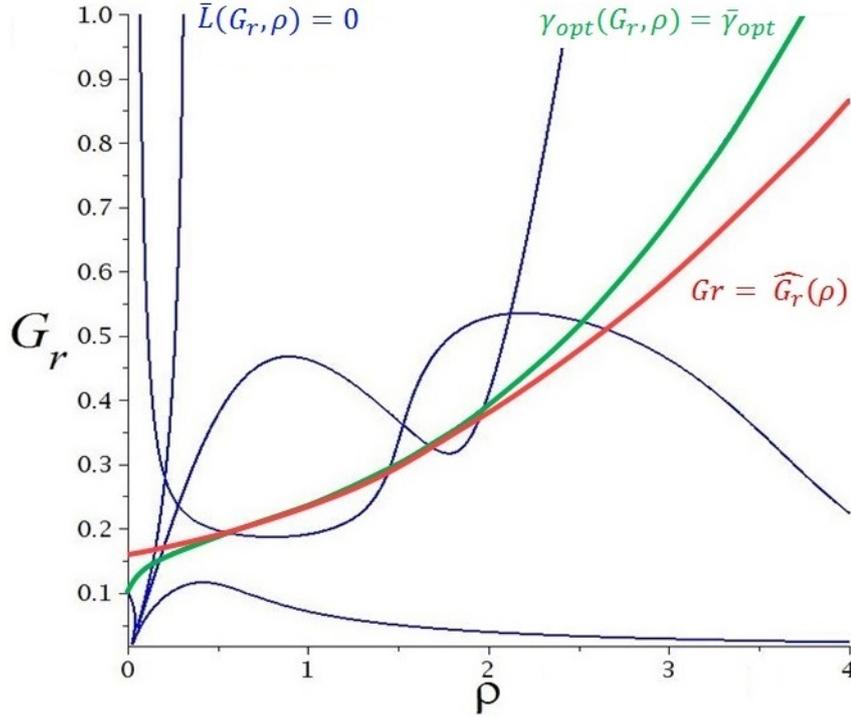


FIGURE 4.16 – Bleu + vert :  $\bar{L}(G_r, \rho) = 0$  – Vert :  $\gamma_{\text{opt}}(G_r, \rho) = \bar{\gamma}_{\text{opt}}$  – Rouge :  $G_r = \widehat{G}_r(\rho)$

isolant ses racines, nous déterminons  $q$  valeurs de  $G_r$ , notées  $G_r^c$ , telles que  $\bar{L}(G_r^c, \rho^c) = 0$ . Nous évaluons ensuite l'expression par radicaux  $\gamma_{\text{opt}}$  en  $(G_r^c, \rho^c)$  pour obtenir :

$$\gamma_{\text{opt}}^c := \gamma_{\text{opt}}(G_r^c, \rho^c).$$

Si  $\gamma_{\text{opt}}^c = \bar{\gamma}_{\text{opt}}$ , alors nous avons :

$$\forall (G_r, \rho) \in \mathcal{C}_i, \quad \bar{L} = 0 \quad \Rightarrow \quad \gamma_{\text{opt}}(G_r, \rho) = \bar{\gamma}_{\text{opt}}.$$

Autrement dit, en tout point de la branche de  $\bar{L}$  de la cellule  $\mathcal{C}_i$ ,  $\gamma_{\text{opt}}$  atteint la valeur souhaitée  $\bar{\gamma}_{\text{opt}}$ . En répétant ce procédé pour chaque cellule  $\mathcal{C}_i$ , nous identifions la branche correspondant à  $\gamma_{\text{opt}}(G_r, \rho) = \bar{\gamma}_{\text{opt}}$  comme la courbe verte de la Figure 4.16.

Afin de résoudre le problème de réglage de pondération, il reste à déterminer une fonction scalaire  $\widehat{G}_r(\rho)$  telle que, pour tout  $\rho \in \mathbb{R}$ , le couple  $(\widehat{G}_r(\rho), \rho)$  appartienne à la branche verte de la Figure 4.16.

Malheureusement, contrairement au cas sans amortissement,  $\bar{L}(G_r, \rho) = 0$  n'est pas résoluble explicitement en  $G_r$ . Par contre, en utilisant un développement de Taylor autour d'un point régulier donné (c'est-à-dire un point n'étant pas sur la variété discriminante de  $L$ ), nous pouvons obtenir une formule explicite de  $G_r$  dépendant de  $\rho$ , dont la courbe associée "reste proche" la courbe verte de la Figure 4.16 au voisinage de ce point.

Par exemple, choisissons le point  $\bar{\rho} := 1$  et considérons un voisinage de ce point. Calculons ensuite un développement en série de Taylor d'ordre 3 de  $\bar{L}(G_r, \rho)$  au point  $\bar{\rho} := 1$ . Nous obtenons alors  $q$  fonctions scalaires  $\widehat{G}_r(\rho)$ , chacune correspondant à l'approximation d'une branche de  $\bar{L}(G_r, \rho)$  au voisinage de  $\bar{\rho} := 1$ . En particulierisant  $\rho = \bar{\rho}$  dans chacune de ces fonctions scalaires, nous obtenons  $q$  valeurs de  $G_r$ , notées  $G_r^q$ . En particulierisant  $(G_r^q, \bar{\rho})$  dans

la formule par radicaux  $\gamma_{\text{opt}}$ , nous identifions quel couple permet d'obtenir :

$$\gamma_{\text{opt}}(G_r^q, \bar{\rho}) = \bar{\gamma}_{\text{opt}}.$$

Nous en déduisons alors la fonction scalaire  $\hat{G}_r$  permettant d'obtenir :

$$\forall \rho \in \mathbb{Q}, \quad \exists (\epsilon_1, \epsilon_2) \in \mathbb{Q}^2, \quad |\rho - \bar{\rho}| < \epsilon_1 \quad \Rightarrow \quad \left| \gamma_{\text{opt}}(\hat{G}_r(\rho), \rho) - \bar{\gamma}_{\text{opt}} \right| < \epsilon_2.$$

La fonction scalaire obtenue est la suivante :

$$\hat{G}_r(\rho) := 0.033 (\rho - 1)^2 + 0.11 \rho + 0.13.$$

Cette fonction est représentée en rouge dans la Figure 4.16. Ainsi, nous pouvons utiliser le paramètre de réglage suivant :

$$\hat{w}(\theta) := m a_2 \hat{G}_r(\rho).$$

En utilisant une telle pondération  $\hat{w}(\theta)$ , nous nous assurons que  $\gamma_{\text{opt}}(\tilde{w}(\theta), \theta)$  reste proche de  $\bar{\gamma}_{\text{opt}}$  autour d'un point de fonctionnement donné. De plus, en notant  $\gamma := \sqrt{\lambda + 1}$ , nous pouvons vérifier que la valeur de  $\gamma_{\text{opt}}$  réelle en traçant  $L(\gamma^2 - 1, \hat{G}_r(\rho), \rho) = 0$ . Voir la Figure 4.17.

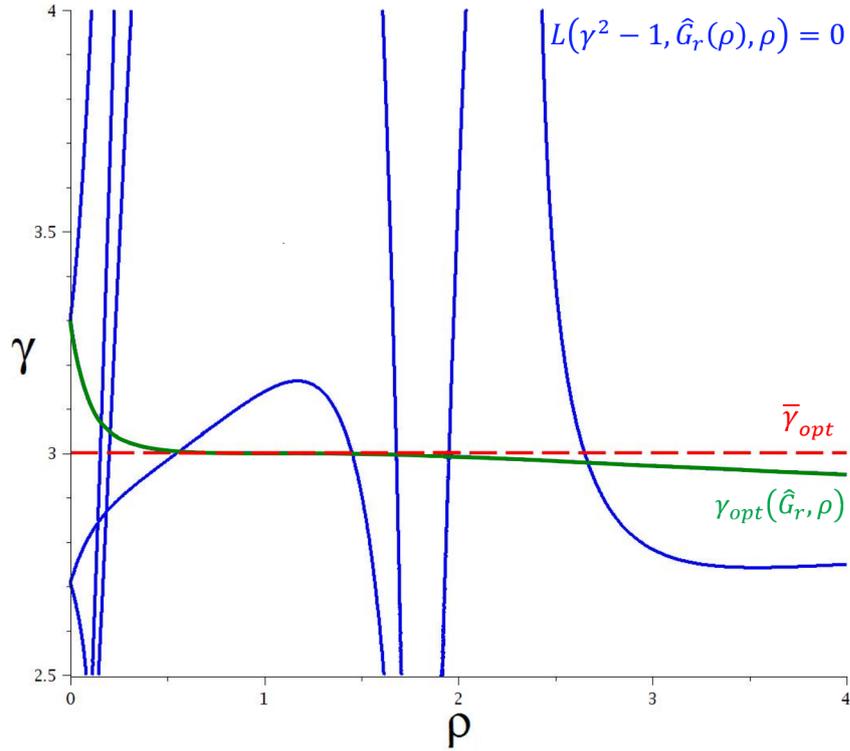


FIGURE 4.17 – Bleu + vert :  $L(\gamma^2 - 1, \hat{G}_r(\rho), \rho) = 0$  – Vert : branche correspondant à  $\gamma_{\text{opt}}(\hat{G}_r(\rho), \rho)$  – Rouge :  $\gamma_{\text{opt}}$  attendu

Disposant de  $\rho$  au voisinage de  $\bar{\rho}$ , nous pouvons alors calculer  $\gamma_{\text{opt}}(\hat{G}_r(\rho), \rho)$  en utilisant nos formules explicites, choisir  $\gamma > \gamma_{\text{opt}}(\hat{G}_r(\rho), \rho)$  et obtenir un correcteur  $C_\gamma(\theta) = K_\gamma(\theta)/\hat{w}(\theta)$  stabilisant  $P$  s'adaptant automatiquement aux variations de  $\rho$  et assurant que  $\gamma_{\text{opt}}(\hat{G}_r(\rho), \rho)$  reste proche de  $\bar{\gamma}_{\text{opt}}$ .

## 4.4 Commande $H_\infty$ adaptative indirecte

L'avantage de disposer de correcteurs explicites est que ceux-ci sont facilement embarquables dans un ordinateur. En effet, l'évaluation de formules par radicaux ne nécessite que des opérations élémentaires telle que l'addition, la multiplication ou le calcul de racines carrées. De telles opérations sont maîtrisées dans les ordinateurs actuels. Nous proposons alors d'utiliser les formules explicites données en Section 4.1 afin de réaliser une loi de commande s'adaptant aux variations des paramètres d'un système.

Soit  $P(\theta)$  le modèle du système réel à asservir, dépendant des paramètres  $\theta$ . Le cadre de nos études sur la correction  $H_\infty$  paramétrique étant limité aux systèmes linéaires invariants, nous supposons que ces *variations paramétriques*  $\theta(t)$  sont lentes (relativement aux dynamiques du système bouclé).

Afin de concevoir un correcteur adaptatif conservant un compromis performance/robustesse quel que soit  $\theta$ , nous utilisons une pondération Loop-Shaping  $w$  et la réglons à l'aide du procédé décrit dans la Section 4.3. Suivant le système  $P$  étudié, nous disposons d'une pondération optimale valable pour tout  $\theta$ , nous parlons alors de pondération optimale globale, ou valable autour d'un point de fonctionnement, et nous parlons de pondération optimale locale. Cette étude se réalise hors ligne à l'aide du modèle du système réel  $P$  et permet de disposer d'une pondération  $w(\theta)$  explicite.

Disposant d'un modèle  $P(\theta)$  du système réel, d'une pondération explicite  $\hat{w}(\theta)$  et d'un correcteur explicite  $K(\hat{\theta})$  stabilisant le système fictif  $G(\theta) := P(\theta)/\hat{w}(\theta)$ , nous proposons de considérer le schéma de contrôle adaptatif décrit en Figure 4.18. Celui-ci requiert un estimateur fournissant un estimé  $\hat{\theta}$  de  $\theta$ . De cette estimée, nous pouvons évaluer des formules explicites d'un correcteur  $K(\hat{\theta})$  stabilisant le système fictif  $G$ . Le réglage de la pondération scalaire  $w$  se faisant hors ligne, nous n'avons plus qu'à évaluer la formule explicite de  $\hat{w}$  dépendant de  $\hat{\theta}$  en ligne pour calculer  $C := K \hat{w}$ . Ce type de procédé est appelé *commande adaptative indirecte* [LLMK11] : les parties identification du modèle et contrôle sont découplées.

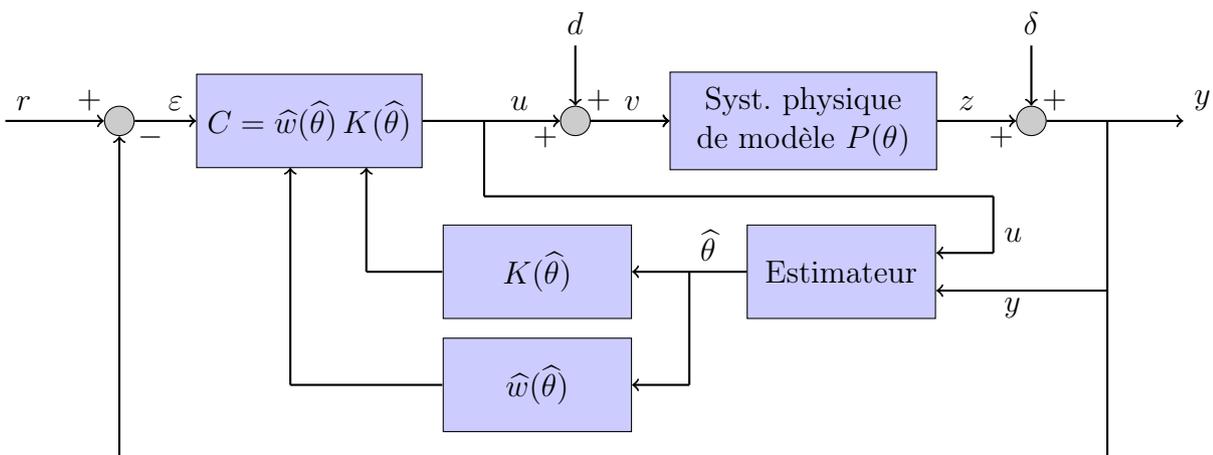


FIGURE 4.18 – Commande adaptative indirecte

Dans les exemples qui suivent, nous considérons que nous disposons d'un *estimateur parfait* permettant d'obtenir en temps réel une mesure exacte  $\hat{\theta}(t) = \theta(t)$  des paramètres du système. La conception d'un estimateur dans ce cadre fera l'objet de travaux futurs. Dans

un premier temps, reprenons l'exemple de la double masse-ressort sans amortissement ainsi que la pondération optimale globale calculée dans la section précédente.

#### 4.4.1 Système double masse-ressort sans amortissement bouclé avec correction adaptative globale

Soit le système double masse-ressort décrit en Section 4.1.3, dont le modèle est :

$$P := \frac{c'_0}{s^2 (s^2 + a_2)}, \quad c'_0 > 0, \quad a_2 > 0.$$

En choisissant la pondération scalaire

$$\begin{aligned} \widehat{w}(a_2, c'_0) &= \frac{4 a_2^2}{197261828370049 c'_0} \left( 6246072 + 3625216 \sqrt{2} - \sqrt{90755864194240 + 67218229682478 \sqrt{2}} \right)^2, \\ &\approx 0.103 \left( \frac{a_2^2}{c'_0} \right), \end{aligned}$$

nous nous assurons que  $\gamma_{\text{opt}}$  associé au système fictif

$$G := \frac{c'_0 \widehat{w}(a_2, c'_0)}{s^2 (s^2 + a_2)}$$

est égal à  $\bar{\gamma}_{\text{opt}} = 3$  quel que soit  $a_2$ . Cette valeur de  $\gamma_{\text{opt}}$  assure de bonnes marges de robustesse en pratique. Alors, en utilisant le Corollaire 2.1 et en choisissant  $\gamma > \gamma_{\text{opt}}$ , nous en déduisons un correcteur  $K$  stabilisant  $G$ . De ce correcteur, en utilisant les propriétés du Loop-Shaping, nous concevons un correcteur  $C$  stabilisant  $P$  qui permet de conserver des propriétés de la boucle ouverte quels que soient  $a_2$  et  $c'_0$ .

Nous effectuons alors une simulation du système  $P$  bouclé par le correcteur adaptatif  $C$ . Cette simulation est représentée dans la Figure 4.19. La courbe bleue représente la référence à atteindre. La courbe rouge représente la sortie du système bouclé avec un correcteur  $C$  fixe calculé au point nominal  $a_2 = 10$  et  $c'_0 = 10$ . La courbe verte représente la sortie du système bouclé avec un correcteur  $C$  recalculé pour chaque valeur de  $a_2$  et  $c'_0$ . Cette simulation est effectuée sur 100 secondes. A  $t = 30s$ , alors que le système a atteint un point d'équilibre, nous changeons brutalement le paramètre  $c'_0$  en le divisant par deux, ainsi que le paramètre  $a_2$  en le multipliant par  $3/4$ . Ceci correspond en pratique à doubler brusquement la masse de la seconde masse. Nous remarquons alors que le système retourne dans un état d'équilibre, avec les *mêmes dynamiques que précédemment*. De plus, nous remarquons que sans adaptation du correcteur, le système bouclé est instable. Ce phénomène n'apparaît pas lorsque le correcteur s'adapte à la variation de  $a_2$ .

#### 4.4.2 Système double masse-ressort avec amortissement bouclé avec correction adaptative locale

Nous proposons maintenant de considérer le même système mais avec un amortissement non nul

$$P := \frac{1}{m} \frac{s + a_2}{s^2 (s^2 + a_3 s + a_2)},$$

et la pondération scalaire locale, valable autour de  $\rho := \frac{a_3}{\sqrt{a_2}} = 1$ , et calculée dans la section précédente :

$$\widehat{w}(a_2, m, \rho) = m a_2 (0.033 (\rho - 1)^2 + 0.11 \rho + 0.13).$$

Nous obtenons alors un correcteur explicite  $C$  stabilisant  $P$ , dépendant de  $\rho$ , et assurant des propriétés de robustesse au voisinage de  $\rho = 1$ . Nous effectuons une simulation du système  $P$  bouclé par le correcteur adaptatif  $C$ . Cette simulation est représentée dans la Figure 4.20. La courbe bleue représente la référence à atteindre. La courbe rouge la sortie du système bouclé avec un correcteur  $C$  fixe calculé au point nominal  $\rho = 1$ . La courbe verte représente la sortie du système bouclé avec un correcteur  $C$  adaptatif au voisinage de  $\rho = 1$ . Cette simulation est effectuée sur 100 secondes. A  $t = 30s$ , alors que le système a atteint un point d'équilibre, nous doublons la masse de la seconde masse. Nous remarquons alors que le système retourne dans un état d'équilibre avec les *mêmes dynamiques que précédemment*. De plus, nous remarquons que notre correcteur adaptatif est nettement plus performant qu'un correcteur nominal.

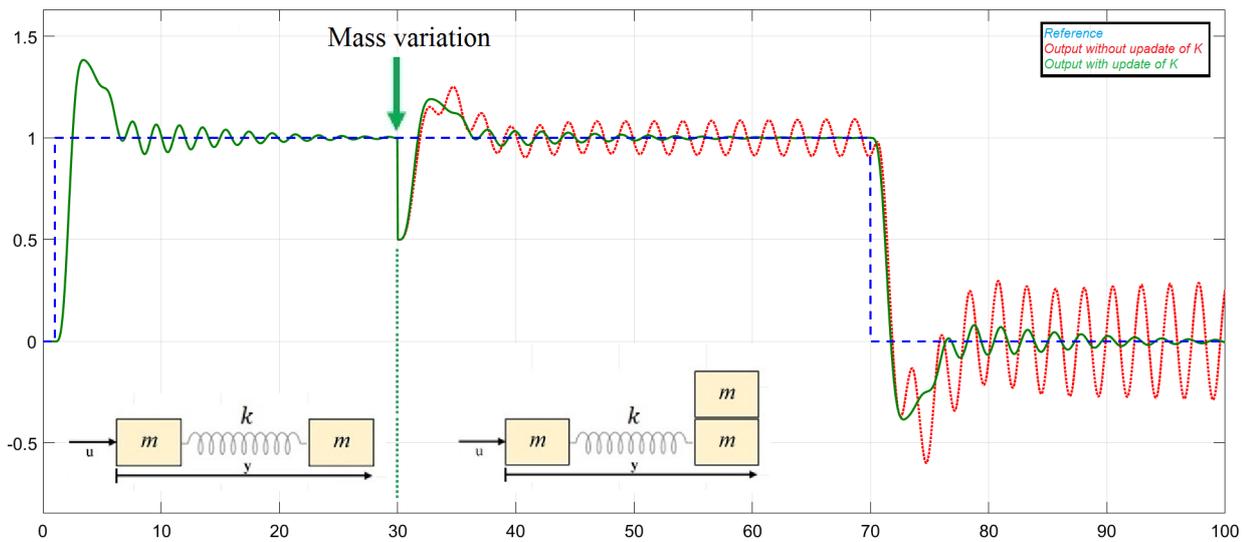


FIGURE 4.19 – Simulation du système sans amortissement bouclé soumis à une variation paramétrique

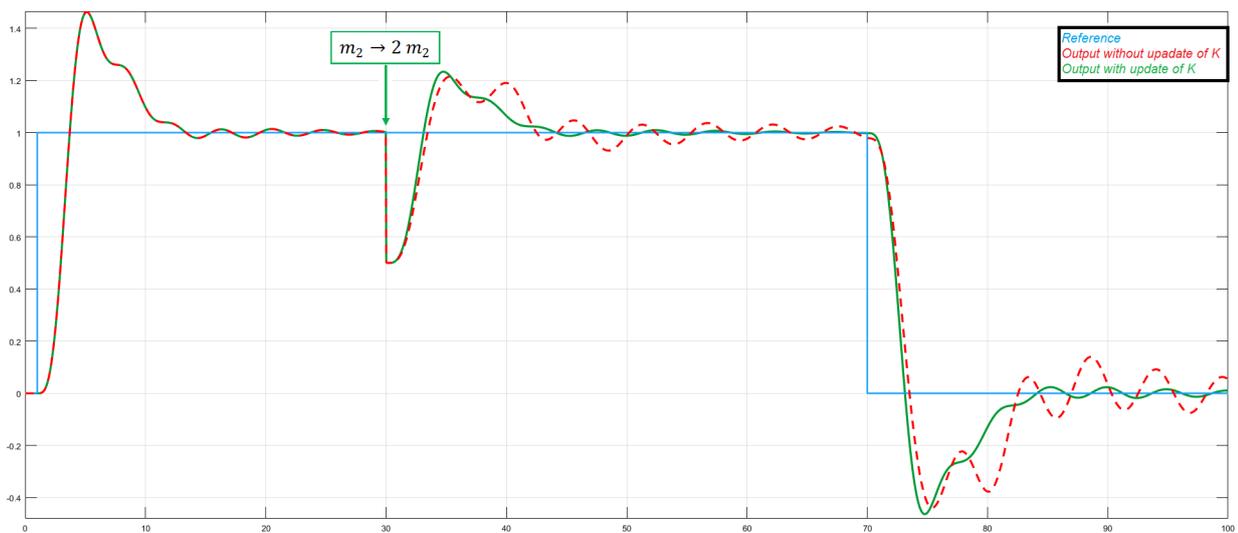


FIGURE 4.20 – Simulation du système avec amortissement bouclé soumis à une variation paramétrique

## 4.5 Conclusion

Ce chapitre a montré comment aborder le problème de commande  $H_\infty$  à paramètres non fixés.

Dans un premier temps, nous avons montré que ce problème pouvait se résoudre entièrement symboliquement pour des systèmes SISO d'ordre inférieur ou égal à 4 à coefficients inconnus quelconques en utilisant des formules par radicaux et en découpant l'espace des paramètres grâce la notion de variété discriminante. Pour des systèmes d'ordre plus élevé, nous avons montré que le calcul formel trouvait aussi son intérêt en proposant des algorithmes de calcul certifié.

Nous avons ensuite montré que l'approche symbolique peut servir pour le choix de pondérations Loop-Shaping. Pour cela, nous avons restreint nos études au cas de pondérations scalaires, et avons montré que l'on pouvait trouver des expressions explicites de celles-ci dépendant des paramètres du système et permettant d'atteindre un critère  $H_\infty$  donné. Suivant la complexité du système, ces réglages de pondération optimaux peuvent se réaliser localement ou globalement.

Plus généralement, ce chapitre a montré que l'approche paramétrique permettait de mieux comprendre la méthode de synthèse  $H_\infty$  que nous utilisons. Disposant de solutions explicites, il est possible de savoir a priori comment va se comporter le système bouclé pour différentes valeurs de paramètres. Cette démarche de compréhension globale du problème de commande permet ensuite à l'ingénieur de gagner du temps dans ses calculs de correcteurs optimaux ou dans l'orientation de l'architecture d'un viseur. Autrement dit, le calcul formel apporte des outils permettant de visualiser des intuitions qu'un ingénieur aurait acquises via son expérience pratique. Ces outils de visualisation pourront être utiles lors de réunions multi-métiers, où une démarche pédagogique et vulgarisatrice est nécessaire afin de mieux se faire comprendre aux autres corps de métier.

En plus de cet impact important sur la compréhension du problème de commande, nous pensons que ces correcteurs paramétriques trouveront leur place dans des perspectives de commande embarquée. En effet, la simplicité des correcteurs obtenus d'un point de vue algorithmique peut-être très attrayante car elle ne nécessite que peu de puissance de calcul. De plus, leur efficacité s'est montrée convaincante en utilisant des estimateurs idéaux. Il reste alors à concevoir des estimateurs réalistes afin de pouvoir comparer ces méthodes à des stratégies de commande robuste non adaptatives. Ces aspects, tant théoriques que pratiques, seront développés dans de futurs travaux.

Le chapitre suivant cherche à étendre ces concepts à certaines classes de systèmes à retard importantes en pratique.



# Chapitre 5

## Extension de la commande $H_\infty$ paramétrique aux systèmes à retard

Dans la Section 4.1, nous avons obtenu des correcteurs sous forme close pour des systèmes linéaires continus SISO invariants dans le temps d'ordre inférieur ou égal à 4, et la Section 4.3 nous a permis d'expliciter des pondérations Loop-Shaping optimales au sens du critère  $H_\infty$ .

Un des désavantages de l'approche paramétrique est que nous sommes rapidement limités par l'ordre des systèmes étudiés. Nous sommes donc amenés à chercher des moyens de complexifier les modélisations des systèmes pouvant être étudiés par notre approche. Dans ce chapitre, nous proposons d'étendre ces résultats à des systèmes à retard.

Ce chapitre utilise des résultats sur la stabilisation de systèmes à retard qui peuvent se décomposer de manière additive [QQ12]. Nous étendons ces résultats au formalisme  $H_\infty$  et proposons une majoration du critère  $H_\infty$  suivant la décomposition choisie. Nous proposons alors un correcteur symbolique sous-optimal au sens du critère  $H_\infty$  défini par (2.6). Enfin, nous illustrons cette approche sur un exemple de modélisation d'un viseur comportant un retard pur. En particulier, nous montrons qu'il est possible de réaliser une commande adaptative indirecte de ce système à retard à l'aide de nos techniques.

Les résultats de ce chapitre sont publiés dans l'article de conférence [RBQQ18].

### Sommaire

---

<b>5.1</b>	<b>Décomposition additive de systèmes à retard et stabilisation</b>	<b>176</b>
<b>5.2</b>	<b>Problème de commande robuste et décomposition additive</b>	<b>178</b>
5.2.1	Liens entre les projecteurs $\Pi(K, G)$ et $\Pi(K_0, G_0)$	178
5.2.2	Un encadrement de $\ \Pi(K, G)\ _\infty$	180
5.2.3	Application à un modèle de viseur gyrostabilisé	182
<b>5.3</b>	<b>Réglage d'un correcteur robuste sous-optimal à l'aide d'une pondération Loop-Shaping</b>	<b>189</b>
5.3.1	Système réel et hypothèses	190
5.3.2	Système fictif et données préliminaires	191
5.3.3	Une procédure de conception de correcteur Loop-Shaping adaptatif stabilisant le $G$	192
5.3.4	Caractéristiques de l'asservissement obtenu	197
<b>5.4</b>	<b>Conclusion</b>	<b>204</b>

---

## 5.1 Décomposition additive de systèmes à retard et stabilisation

Soit  $F \in RH_\infty$  tel que  $F(0) \neq 0$ . Soit  $\tau > 0$  et  $G$  un système à retard SISO de la forme suivante :

$$G = \frac{F(s)}{s} e^{-\tau s}. \quad (5.1)$$

Nous notons le système  $G$  sans retard par

$$G_0 := \frac{F(s)}{s}, \quad (5.2)$$

c'est-à-dire  $G$  dans lequel nous avons particularisé  $\tau = 0$ . Notons aussi :

$$\Delta := \left( \frac{1 - e^{-\tau s}}{s} \right) F(s). \quad (5.3)$$

Remarquons alors que nous avons :

$$\frac{1 - e^{-\tau s}}{s} = \tau + O(s^2).$$

$\frac{1 - e^{-\tau s}}{s}$  n'a donc aucun pôle dans  $\mathbb{C}_+$ . De plus, sa norme  $H_\infty$  est bornée par  $\tau$ . Par conséquent, nous avons :

$$\Delta \in H_\infty(\mathbb{C}_+).$$

Nous pouvons alors écrire  $G$  comme la somme de deux systèmes :

$$G = G_0 - \Delta. \quad (5.4)$$

Nous avons le théorème suivant.

**Théorème 5.1.** [QQ12] Soit  $F \in RH_\infty$  tel que  $F(0) \neq 0$ . Soient  $G_0 := F(s)/s \in \mathbb{R}(s)$ ,  $G = G_0 e^{-\tau s}$  et  $\Delta := \left( \frac{1 - e^{-\tau s}}{s} \right) F \in H_\infty(\mathbb{C}_+)$ . Enfin, soit  $K_0 \in \mathbb{R}(s)$  un correcteur stabilisant  $G_0$ . Alors, le correcteur suivant stabilise  $P$  de manière interne :

$$K := \frac{K_0}{1 + \Delta K_0}. \quad (5.5)$$

Le Théorème 5.1 donne un correcteur  $K$  stabilisant le système avec retard  $G$  en utilisant un correcteur  $K_0$  stabilisant le système sans retard  $G_0$ . Un tel résultat peut se décrire par le schéma de la Figure 5.1.

**Remarque 5.1.** Si  $G_0$  est d'ordre inférieur ou égal à 4, alors nous pouvons concevoir un correcteur stabilisant  $K$  dépendant explicitement des paramètres de  $G$ .

**Remarque 5.2.** Si  $\Delta = 0$ , alors nous obtenons de nouveau le schéma de contrôle classique décrit dans la Figure 5.2.

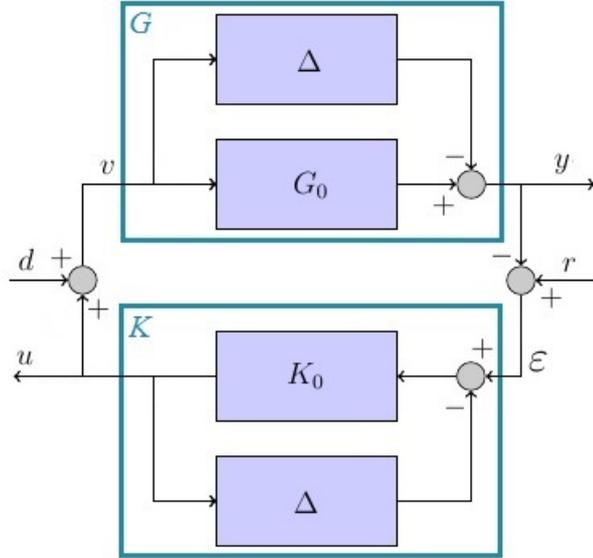
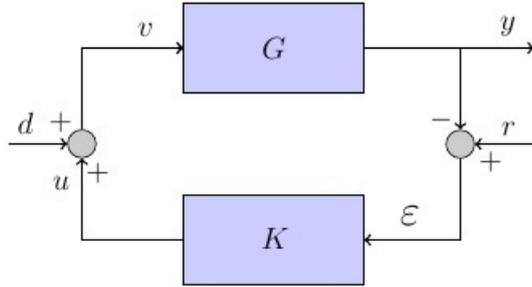


FIGURE 5.1 – Schéma d'asservissement d'un système avec retard


 FIGURE 5.2 – Schéma d'asservissement avec  $\Delta = 0$ 

L'idée est alors de combiner ce théorème avec nos résultats obtenus sur la commande  $H_\infty$  paramétrique pour des systèmes de dimension finie d'ordre faible. En particulier, si  $G_0$  est de degré inférieur ou égal à 4, alors, dans la Section 4.1, nous avons explicité des correcteurs  $H_\infty$  paramétriques dépendant des paramètres inconnus de  $G_0$  satisfaisant le PCR. Plus précisément, en définissant le projecteur  $\Pi$  suivant

$$\Pi(K_0, G_0) := \begin{pmatrix} S_0 & S_0 G_0 \\ K_0 S_0 & K_0 S_0 G_0 \end{pmatrix}, \quad S_0 := (1 + K_0 G_0)^{-1}, \quad (5.6)$$

ces correcteurs satisfont alors (voir Définition 2.7 du Chapitre 2)

$$\|\Pi(K_0, G_0)\|_\infty \leq \gamma_{\text{opt}}, \quad (5.7)$$

où  $\gamma_{\text{opt}}$  est déterminé explicitement en fonction des paramètres de  $G_0$ . Disposant d'un correcteur  $H_\infty$  explicite  $K_0$  stabilisant le système  $G_0$  d'ordre inférieur ou égal à 4, nous déduisons alors des correcteurs  $H_\infty$  explicites  $K$  stabilisant le système  $G$ .

Cependant, même si  $K_0$  stabilise  $G_0$  et satisfait le PCR, c'est-à-dire vérifie (5.7), alors nous ne pouvons rien dire de  $\|\Pi(K, G)\|_\infty$ . Nous pouvons supposer que celle-ci sera plus grande que  $\gamma_{\text{opt}}$ , un système à retard étant plus "complexe" à stabiliser qu'un système sans retard du fait du caractère de dimension infinie d'un système à retard. La section suivante cherche à obtenir une estimation sur cette dernière norme.



nous avons alors :

$$D^{-1} = \begin{pmatrix} 1 & -\Delta \\ 0 & 1 \end{pmatrix}.$$

Enfin, comme  $\Delta \in H_\infty(\mathbb{C}_+)$ , alors  $D \in H_\infty(\mathbb{C}_+)^{2 \times 2}$  et  $D^{-1} \in H_\infty(\mathbb{C}_+)^{2 \times 2}$ . Notons aussi que  $\Pi(K_0, G_0) \in H_\infty(\mathbb{C}_+)^{2 \times 2}$  puisque  $K_0$  stabilise  $G_0$ . Ainsi, nous obtenons  $\Pi(K, G) \in H_\infty(\mathbb{C}_+)^{2 \times 2}$  car défini comme un produit d'éléments de  $H_\infty(\mathbb{C}_+)$ .  $\square$

**Proposition 5.1.** La dérivée du projecteur  $\Pi$  par rapport à  $\Delta$  a alors la forme particulière suivante :

$$\begin{aligned} \partial \Pi &:= \lim_{\Delta \rightarrow 0} \frac{\Pi(K, G) - \Pi(K_0, G_0)}{\Delta} = L \Pi(K_0, G_0) - \Pi(K_0, G_0) L \\ &= [L, \Pi(K_0, G_0)], \end{aligned} \quad (5.10)$$

où  $[A, B] = AB - BA$  est le *commutateur* de  $A$  et  $B$ . Notons que

$$\text{ad}_L : \Pi(K_0, G_0) \longmapsto [L, \Pi(K_0, G_0)]$$

est une *dérivée intérieure* de l'anneau de Lie  $H_\infty(\mathbb{C}_+)^{2 \times 2}$ . Cette forme particulière sera exploitée dans de futurs travaux.

*Démonstration.* Les équations (5.9) permettent d'écrire  $\Pi$  sous la forme suivante :

$$\Pi(K, G) = \begin{pmatrix} S & GS \\ KS & GKS \end{pmatrix} = \Pi(K_0, G_0) + \Delta \begin{pmatrix} S_0 K_0 & T_0 - S_0 \\ 0 & -S_0 K_0 \end{pmatrix} - \Delta^2 \begin{pmatrix} 0 & S_0 K_0 \\ 0 & 0 \end{pmatrix}.$$

En remarquant que (voir (5.9))

$$T_0 - S_0 = T_0 - \Delta S_0 K_0 - (S_0 - \Delta S_0 K_0) = T - S + 2 \Delta K S,$$

nous obtenons alors :

$$\Pi(K, G) = \Pi(K_0, G_0) + \Delta \begin{pmatrix} SK & T - S \\ 0 & -SK \end{pmatrix} + \Delta^2 \begin{pmatrix} 0 & SK \\ 0 & 0 \end{pmatrix}.$$

De manière équivalente, cette dernière égalité se réécrit sous la forme suivante :

$$\Pi(K_0, G_0) = \Pi(K, G) - \Delta \begin{pmatrix} SK & T - S \\ 0 & -SK \end{pmatrix} - \Delta^2 \begin{pmatrix} 0 & SK \\ 0 & 0 \end{pmatrix}.$$

En définissant  $L := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , nous avons alors

$$\begin{cases} D = I_2 + \Delta L, \\ D^{-1} = I_2 - \Delta L, \end{cases}$$

où  $I_2$  désigne la matrice identité de taille  $2 \times 2$ . Le Théorème 5.2 permet alors d'obtenir

$$\Pi(K, G) = (I_2 + \Delta L) \Pi(K_0, G_0) (I_2 - \Delta L),$$

ce que l'on peut réécrire sous la forme de :

$$\Pi(K, G) = \Pi(K_0, G_0) + \Delta (L \Pi(K_0, G_0) - \Pi(K_0, G_0) L) - \Delta^2 L \Pi(K_0, G_0) L.$$

Nous en déduisons alors l'expression (5.10).  $\square$

### 5.2.2 Un encadrement de $\|\Pi(K, G)\|_\infty$

Nous avons montré que  $\Pi(K, G) \in H_\infty(\mathbb{C}_+)^{2 \times 2}$ . Ainsi, sa norme  $H_\infty$  est finie. Le théorème suivant montre que nous pouvons en calculer un majorant.

**Théorème 5.3.** [RBQQ18] Avec les notations précédentes, la matrice  $D := \begin{pmatrix} 1 & \Delta \\ 0 & 1 \end{pmatrix} \in H_\infty(\mathbb{C}_+)^{2 \times 2}$  vérifie que :

$$\|D^{-1}\|_\infty = \|D\|_\infty.$$

De plus, nous avons :

$$\|D\|_\infty = \sqrt{1 + \frac{\|\Delta\|_\infty^2}{2}} + \sqrt{\left(1 + \frac{\|\Delta\|_\infty^2}{2}\right)^2 - 1}. \quad (5.11)$$

Finalement, l'inégalité suivante est vérifiée :

$$\frac{\|\Pi(K_0, G_0)\|_\infty}{\|D\|_\infty^2} \leq \|\Pi(K, G)\|_\infty \leq \|D\|_\infty^2 \|\Pi(K_0, G_0)\|_\infty. \quad (5.12)$$

*Démonstration.* En passant à la norme  $H_\infty$ , le Théorème 5.2 implique l'égalité suivante :

$$\|\Pi(K, G)\|_\infty = \|D \Pi(K_0, G_0) D^{-1}\|_\infty.$$

Le Théorème 5.2 donne aussi  $\Pi(K_0, G_0) = D^{-1} \Pi(K, G) D$ , soit en passant à la norme  $H_\infty$  :

$$\|\Pi(K_0, G_0)\|_\infty = \|D^{-1} \Pi(K, G) D\|_\infty.$$

En utilisant les propriétés classiques de la norme  $H_\infty$  (voir Annexe B ou [Vin00]), nous obtenons alors :

$$\frac{\|\Pi(K_0, G_0)\|_\infty}{\|D\|_\infty \|D^{-1}\|_\infty} \leq \|\Pi(K, G)\|_\infty \leq \|D\|_\infty \|\Pi(K_0, G_0)\|_\infty \|D^{-1}\|_\infty. \quad (5.13)$$

Nous désirons maintenant calculer la norme  $H_\infty$  de  $D$ . Pour cela, calculons le polynôme caractéristique de  $D(i\omega)D^*(i\omega)$  :

$$\begin{aligned} h_D &:= \det(\lambda I - D(i\omega)D^*(i\omega)) \\ &= \det \left( \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} - \begin{pmatrix} 1 & \Delta(i\omega) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \Delta^*(i\omega) & 1 \end{pmatrix} \right) \\ &= \begin{vmatrix} \lambda - 1 - \Delta(i\omega) \Delta^*(i\omega) & -\Delta(i\omega) \\ -\Delta^*(i\omega) & \lambda - 1 \end{vmatrix} \\ &= (\lambda - 1 - \Delta(i\omega) \Delta^*(i\omega)) (\lambda - 1) - \Delta(i\omega) \Delta^*(i\omega) \\ &= (\lambda - 1 - |\Delta(i\omega)|^2) (\lambda - 1) - |\Delta(i\omega)|^2 \\ &= \lambda^2 - 2\lambda + 1 - |\Delta(i\omega)|^2 \lambda \\ &= \lambda^2 - (2 + |\Delta(i\omega)|^2) \lambda + 1. \end{aligned} \quad (5.14)$$

Notons que le polynôme caractéristique de  $D^{-1}(i\omega)D^{-1*}(i\omega)$  est égal à  $h_D$ . Ceci se vérifie immédiatement en remplaçant  $\Delta$  par  $-\Delta$  dans  $h_D$ . Ainsi, nous avons :

$$\|D^{-1}\|_\infty = \|D\|_\infty. \quad (5.15)$$

Par ailleurs, pour  $\epsilon \in \{-1, 1\}$ , les racines de  $h_D$  en  $\lambda$  sont données par :

$$\begin{aligned}\lambda_\epsilon(\omega) &:= \frac{(2 + |\Delta(i\omega)|^2) + \epsilon \sqrt{|\Delta(i\omega)|^4 + 4|\Delta(i\omega)|^2}}{2} \\ &= 1 + \frac{|\Delta(i\omega)|^2}{2} + \epsilon \sqrt{\left(1 + \frac{|\Delta(i\omega)|^2}{2}\right)^2 - 1}.\end{aligned}$$

Notons que pour toute valeur de  $\omega \in \mathbb{R}$ ,  $|\Delta(i\omega)|^2 \geq 0$ . Ainsi, nous avons :

$$\forall \omega \in \mathbb{R}, \quad \lambda_1(\omega) \geq \lambda_{-1}(\omega).$$

Par conséquent, nous obtenons :

$$\|D\|_\infty = \sup_{\omega \in \mathbb{R}} \sqrt{\lambda_1(\omega)}. \quad (5.16)$$

Notons  $\omega_m$  une des fréquences pour laquelle  $|\Delta(i\omega)|$  est maximal. Nous avons alors :

$$\|\Delta\|_\infty = |\Delta(i\omega_m)|. \quad (5.17)$$

En utilisant (5.16) et (5.17), nous déduisons :

$$\|D\|_\infty = \sqrt{1 + \frac{\|\Delta\|_\infty}{2}} + \sqrt{\left(1 + \frac{\|\Delta\|_\infty}{2}\right)^2 - 1}.$$

□

**Remarque 5.3.** Supposons  $\|\Delta\|_\infty = 1$  (par exemple avec  $\tau = 1$  et  $F = 1$ ). Alors, nous avons

$$\|D\|_\infty = \sqrt{1 + \frac{1 + \sqrt{5}}{2}} = \sqrt{\frac{3 + \sqrt{5}}{2}} = \frac{1 + \sqrt{5}}{2},$$

c'est-à-dire :

$$\|\Pi(K, G)\|_\infty \leq \left(\frac{1 + \sqrt{5}}{2}\right)^2 \|\Pi(K_0, G_0)\|_\infty = \frac{3 + \sqrt{5}}{2} \|\Pi(K_0, G_0)\|_\infty.$$

Nous voyons apparaître le *nombre d'or*  $\varphi = \frac{1 + \sqrt{5}}{2}$  !

**Remarque 5.4.** Un développement de Taylor de  $\|D\|_\infty^2$  donné par (5.11) pour  $\|\Delta\|_\infty$  en 0 donne :

$$\|\Pi(K, G)\|_\infty \leq \left(1 + \|\Delta\|_\infty + \frac{1}{2} \|\Delta\|_\infty^2 + \frac{1}{8} \|\Delta\|_\infty^3 + O\left(\|\Delta\|_\infty^5\right)\right) \|\Pi(K_0, G_0)\|_\infty.$$

Un tel développement limité pourra être, par exemple, utilisé dans un calculateur embarqué.

Par ailleurs, dans la Figure 5.3, nous avons représenté l'évolution de  $\|D\|_\infty^2$  et  $\|D\|_\infty^{-2}$ . Ainsi, nous remarquons que lorsque  $\|\Delta\|_\infty$  est très grande, alors  $\|D\|_\infty^2$  est grande et  $\|D\|_\infty^{-2}$  est proche de 0. Dans ce cas, l'inégalité (5.13) devient très large. De tels cas apparaissent lorsque  $G_0$  comporte des modes mécaniques très résonants et/ou lorsque le gain de  $\Delta$ , c'est-à-dire de  $F$ , est très grand.

La borne supérieure (5.13) de  $\|\Pi(K, G)\|_\infty$  peut être grande simplement parce que  $\|\Pi(K, G)\|_\infty$  est grande, comme elle peut le devenir parce que  $\|\Delta\|_\infty$  est grande sans que pour autant  $\|\Pi(K, G)\|_\infty$  le soit. Autrement dit, il peut y avoir des simplifications dans  $\Pi(K, G)$  telles que sa norme soit faible bien que  $\|\Delta\|_\infty$  soit grand, et ceci n'est pas vérifiable en utilisant des inégalités sur la norme  $H_\infty$ . En pratique, nous observons que cette borne n'est pas proche de  $\|\Pi(K, G)\|_\infty$  lorsque  $\Delta$  a son maximum ailleurs qu'en zéro.

Nous proposons d'appliquer ces résultats sur un exemple pratique. Ceci nous permettra d'illustrer nos propos concernant la non-viabilité de cette borne pour un système résonnant.

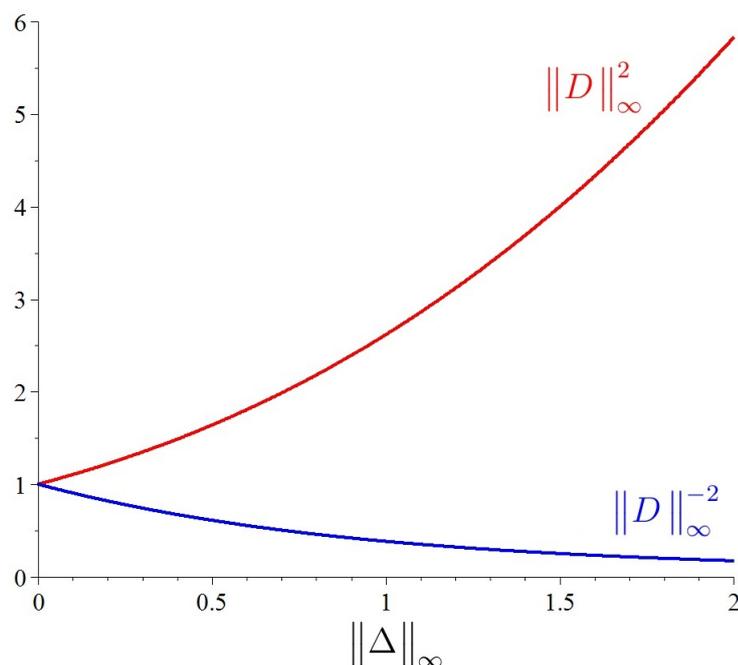


FIGURE 5.3 – Évolution de  $\|D\|_\infty^2$  et  $\|D\|_\infty^{-2}$  en fonction de  $\|\Delta\|_\infty$

### 5.2.3 Application à un modèle de viseur gyrostabilisé

Nous proposons d'appliquer l'approche précédente à un modèle de viseur gyrostabilisé d'ordre 4 avec un retard : le système  $G_{4,r}$  donné dans la Table G.3 est défini par

$$G := G_0 e^{-\tau s},$$

où :

$$G_0 := \frac{g_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \frac{1}{Ts + 1}, \quad T > 0.$$

En notant  $g = g_0 \left( \frac{\omega_0}{\omega_1} \right)^2$ , nous pouvons réécrire la fonction de transfert sous la forme suivante :

$$G = \frac{g}{s} \left( \frac{\frac{s^2}{\omega_0^2} + 2\xi\frac{s}{\omega_0} + 1}{\frac{s^2}{\omega_1^2} + 2\xi\frac{s}{\omega_1} + 1} \right) \frac{1}{Ts + 1} e^{-\tau s}.$$

Comme précédemment, nous adoptons les notations suivantes :

$$\begin{cases} F := g \left( \frac{\frac{s^2}{\omega_0^2} + 2\xi \frac{s}{\omega_0} + 1}{\frac{s^2}{\omega_1^2} + 2\xi \frac{s}{\omega_1} + 1} \right) \frac{1}{Ts + 1} \in RH_\infty, \\ G_0 := \frac{F(s)}{s} \in \mathbb{R}(s), \\ \Delta := \left( \frac{1 - e^{-\tau s}}{s} \right) F(s) \in H_\infty(\mathbb{C}_+). \end{cases}$$

Nous supposons que les paramètres physiques satisfont les propriétés suivantes :

$$0 < \xi \leq 1, \quad T > 0, \quad \tau > 0, \quad \omega_0 > 0, \quad \omega_1 > \omega_0. \quad (5.18)$$

Nous avons le lemme suivant.

**Lemme 5.1.** La norme  $H_\infty$  de  $\Delta$  est bornée par :

$$\|\Delta\|_\infty \leq \tau \|F\|_\infty. \quad (5.19)$$

*Démonstration.* Soit  $\omega \in \mathbb{R}_{+*}$ . Nous avons :

$$|\Delta(i\omega)| = \left| \frac{1 - e^{-\tau i\omega}}{i\omega} \right| |F(i\omega)|.$$

En désignant par  $\text{sin}_c$  la fonction *sinus cardinal*, c'est-à-dire  $\text{sin}_c(x) = \sin(x)/x$  pour tout  $x \in \mathbb{R}$ , et en utilisant l'identité classique  $2 \sin^2(a) = 1 - \cos(2a)$ , nous avons alors :

$$\begin{aligned} \left| \frac{1 - e^{-\tau i\omega}}{i\omega} \right| &= \frac{1}{\omega} |1 - (\cos(\tau\omega) - i \sin(\tau\omega))| \\ &= \frac{1}{\omega} \sqrt{(1 - \cos(\tau\omega))^2 + \sin(\tau\omega)^2} \\ &= \frac{1}{\omega} \sqrt{1 - 2 \cos(\tau\omega) + \cos(\tau\omega)^2 + \sin(\tau\omega)^2} \\ &= \frac{\sqrt{2}}{\omega} \sqrt{1 - \cos(\tau\omega)} \\ &= \frac{\sqrt{2}}{\omega} \sqrt{2 \sin^2\left(\frac{\tau\omega}{2}\right)} \\ &= \frac{2}{\omega} \left| \sin\left(\frac{\tau\omega}{2}\right) \right| \\ &= \tau \left| \text{sin}_c\left(\frac{\tau\omega}{2}\right) \right|. \end{aligned}$$

Comme la fonction  $\text{sin}_c$  admet 1 pour maximum global à l'origine, nous obtenons alors

$$\|\Delta\|_\infty \leq \tau \|F\|_\infty,$$

ce qui conclut la preuve. □

Le Lemme 5.1 permet alors de déduire un majorant suivant de  $\|D\|_\infty^2$  :

$$\|D\|_\infty^2 \leq 1 + \frac{\tau^2}{2} \|F\|_\infty^2 + \sqrt{\left(1 + \frac{\tau^2}{2} \|F\|_\infty^2\right)^2 - 1}.$$

Nous en déduisons un majorant de  $\|\Pi(K, G)\|_\infty$  dépendant de  $\|F\|_\infty$  :

$$\|\Pi(K, G)\|_\infty \leq \left( 1 + \frac{\tau^2}{2} \|F\|_\infty^2 + \sqrt{\left( 1 + \frac{\tau^2}{2} \|F\|_\infty^2 \right)^2 - 1} \right) \|\Pi(K_0, G_0)\|_\infty. \quad (5.20)$$

Nous proposons maintenant de trouver un majorant explicite de  $\|F\|_\infty$  en fonction des paramètres de  $G$ . Pour cela, notons :

$$R := \frac{\left(\frac{s}{\omega_0}\right)^2 + 2\xi\left(\frac{s}{\omega_0}\right) + 1}{\left(\frac{s}{\omega_1}\right)^2 + 2\xi\left(\frac{s}{\omega_1}\right) + 1}.$$

Sachant que

$$\forall T > 0, \quad \left\| \frac{1}{1 + T s} \right\|_\infty = 1,$$

alors :

$$\|F\|_\infty \leq g \|R\|_\infty \left\| \frac{1}{1 + \tau_1 s} \right\|_\infty \leq g \|R\|_\infty. \quad (5.21)$$

De (5.20) et (5.21), nous en déduisons un majorant de  $\|\Pi(K, G)\|_\infty$  dépendant de  $\|R\|_\infty$  :

$$\|\Pi(K, G)\|_\infty \leq \left( 1 + \frac{g^2 \tau^2}{2} \|R\|_\infty^2 + \sqrt{\left( 1 + \frac{g^2 \tau^2}{2} \|R\|_\infty^2 \right)^2 - 1} \right) \|\Pi(K_0, G_0)\|_\infty. \quad (5.22)$$

Il reste alors à calculer la norme  $H_\infty$  de  $R$ . Celle-ci est donnée dans la proposition suivante.

**Proposition 5.2.** Soient  $\omega_0 \in \mathbb{R}_{+*}$ ,  $\omega_1 \in \mathbb{R}_{+*}$ ,  $\omega_0 \neq \omega_1$ ,  $0 < \xi \leq 1$  et :

$$R := \frac{\left(\frac{s}{\omega_0}\right)^2 + 2\xi\left(\frac{s}{\omega_0}\right) + 1}{\left(\frac{s}{\omega_1}\right)^2 + 2\xi\left(\frac{s}{\omega_1}\right) + 1}.$$

Soient  $r := \frac{\omega_1}{\omega_0}$  et :

$$\delta := (r^2 + 1)^2 - 4r^2(2\xi^2 - 1)^2.$$

Alors, la norme  $H_\infty$  de  $R$  est donnée par :

$$\|R\|_\infty = \begin{cases} 1 & \text{si } \xi \geq \frac{\sqrt{2}}{2}, r < 1, \\ r^2 & \text{si } \xi \geq \frac{\sqrt{2}}{2}, r > 1, \\ r \sqrt{\frac{1 - r^2 + \sqrt{\delta}}{r^2 - 1 + \sqrt{\delta}}} & \text{si } \xi < \frac{\sqrt{2}}{2}, r < 1, \\ r \sqrt{\frac{r^2 - 1 + \sqrt{\delta}}{1 - r^2 + \sqrt{\delta}}} & \text{si } \xi < \frac{\sqrt{2}}{2}, r > 1. \end{cases} \quad (5.23)$$

*Démonstration.* Soit  $\omega \in \mathbb{R}$ . Rappelons que la norme  $H_\infty$  de  $R$  est définie par :

$$\|R\|_\infty = \sup_{\omega \in \mathbb{R}} |R(i\omega)|.$$

Soient  $\Omega := \frac{\omega}{\omega_1} \in \mathbb{R}$  et  $r := \frac{\omega_1}{\omega_0}$ . Alors, en appliquant ces changements de variable à  $R(i\omega)$ , nous obtenons :

$$\bar{R}(i\Omega) := \frac{1 - r^2 \Omega^2 + 2\xi r \Omega i}{1 - \Omega^2 + 2\xi \Omega i}.$$

Le changement de variable  $\omega \mapsto \Omega$  est usuellement appelé *normalisation en fréquence*. Lorsque nous effectuons un tel changement de variable, la norme  $H_\infty$  reste inchangée de par la bijectivité de  $\omega \mapsto \Omega$ . Nous préférons alors travailler avec les variables réduites  $\Omega$  et  $r$ , afin de réduire le nombre de paramètres de 3 à 2. De plus, notons que :

$$|\bar{R}(i\Omega)| = \sqrt{\frac{(1 - r^2 \Omega^2)^2 + 4\xi^2 r^2 \Omega^2}{(1 - \Omega^2)^2 + 4\xi^2 \Omega^2}} = |\bar{R}(-i\Omega)|.$$

Ainsi, nous avons :

$$\|R\|_\infty = \max_{\Omega \in \mathbb{R}} |\bar{R}(i\Omega)| = \max_{\Omega \in \mathbb{R}_+} |\bar{R}(i\Omega)|.$$

Dans la suite, nous considèrerons alors  $\Omega \in \mathbb{R}_+$ . Cherchons les pulsations  $\Omega$  où  $|R(i\Omega)|$  admet un extremum. Notons :

$$P := \Omega (r^2(2\xi^2 - 1)\Omega^4 + (r^2 + 1)\Omega^2 + 2\xi^2 - 1).$$

L'expression de la dérivée de  $|R(i\Omega)|$  en  $\Omega$  donne :

$$\frac{\partial |\bar{R}(i\Omega)|^2}{\partial \Omega} := \frac{4(r-1)(r+1)P(\Omega, r, \xi)}{((1-\Omega)^2 + 4\xi^2 \Omega^2)^2}.$$

Maintenant,  $\xi$  étant supposé strictement positif, le dénominateur de  $\partial |\bar{R}(i\Omega)|^2 / \partial \Omega$  ne peut s'annuler et est strictement positif. Par hypothèse  $r = \frac{\omega_1}{\omega_0} > 1$  (voir (5.18)) donc  $|\bar{R}(i\Omega)|^2$  est extrémal si et seulement si  $P$  s'annule.  $P$  est défini par le produit de  $\Omega$  et d'un polynôme de degré 2 en  $\Omega^2$  dont les racines peuvent s'exprimer par radicaux (voir Annexe D). Soit  $\varepsilon \in \{-1, 1\}$  et notons :

$$\delta(r, \xi) := (r^2 + 1)^2 - 4r^2(2\xi^2 - 1)^2.$$

Notons que  $\delta$  représente le discriminant de  $P/\Omega$  en  $\Omega^2$ . En calculant les extremums de  $\delta(r, \xi)$  et en utilisant les hypothèses faites sur les paramètres (voir (5.18)), nous obtenons les inégalités suivantes :

$$\begin{cases} r > 1, \\ \delta(r, \xi) > 0. \end{cases}$$

Notons alors :

$$\Omega_\varepsilon := \sqrt{\frac{-(r^2 + 1) + \varepsilon \sqrt{\delta(r, \xi)}}{2r^2(2\xi^2 - 1)}}.$$

Alors les racines de  $P$  sont :

$$E_\Omega := \{\Omega \mid P(\Omega) = 0\} = \{0, \Omega_{-1}, \Omega_1, -\Omega_{-1}, -\Omega_1\}.$$

Les racines de  $P$  définissent les fréquences  $\Omega$  où  $|\bar{R}(i\Omega)|$  admet un extremum. Il reste à déterminer quel  $\Omega \in E_\Omega$  permet de retrouver le maximum de  $|\bar{R}(i\Omega)|$ . L'expression de ce maximum peut changer lorsque les paramètres varient. Par conséquent, considérons le discriminant de  $P$  en  $\Omega$  :

$$\Delta_\Omega := 4r(2\xi^2 - 1)\delta(r, \xi)^2.$$

Lorsque  $\Delta_\Omega = 0$ , l'expression par radicaux de la pulsation correspondant au maximum de  $|\bar{R}(i\Omega)|$  peut changer. Voir l'Annexe D. Par conséquent, nous avons l'équivalence suivante :

$$\Delta_\Omega = 0 \Leftrightarrow 2\xi^2 - 1 = 0 \Leftrightarrow \xi = \xi_c := \frac{\sqrt{2}}{2}.$$

Disposant de ces limites, nous pouvons séparer l'espace des paramètres en cellules dans lesquelles le nombre de racines réelles de  $P$  est constant, comme le montre la Figure 5.4. L'espace des paramètres  $(\xi, r)$ , limité pour  $\xi \in ]0, 1]$  et  $r > 0, r \neq 1$ , est divisé en deux cellules, une rouge et une bleue, délimitées par la droite  $\xi = \xi_c$ , seule variété sur laquelle le discriminant s'annule.

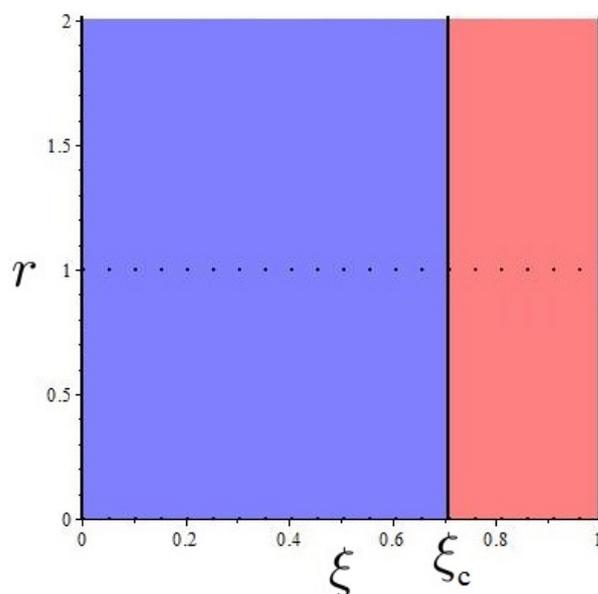


FIGURE 5.4 – Décomposition en cellules de l'espace des paramètres  $\{\xi, r\}$

En prenant un point particulier dans chaque cellule et en évaluant les formes closes  $\Omega_\varepsilon$ , on montre que :

- Si  $\xi > \xi_c$  (cellule rouge), alors 0 est la seule racine réelle de  $P$ .
- Si  $\xi < \xi_c$  (cellule bleue), alors  $P$  a 5 racines réelles et  $\Omega_1 > \Omega_{-1} > 0 > -\Omega_{-1} > -\Omega_1$ .

Nous remarquons aussi que :

$$\begin{cases} \lim_{\Omega \rightarrow 0} |R(i\Omega)| = 1, \\ \lim_{\Omega \rightarrow +\infty} |R(i\Omega)| = r^2. \end{cases} \quad (5.24)$$

Suivants ces deux cas distincts, nous déduisons alors les résultats suivants sur  $\|R\|_\infty$ .

• **Cas 1 :  $\xi \geq \xi_c$  (cellule rouge).**

$|\bar{R}(i\Omega)|$  admet un unique extremum en  $\Omega = 0$ . Ainsi,  $|\bar{R}(i\Omega)|$  est monotone sur  $[0, +\infty[$ . Par conséquent, en utilisant (5.24), nous avons :

$$\forall \xi \geq \xi_c, \quad \|F\|_\infty = \begin{cases} 1 & \text{si } r < 1, \\ r^2 & \text{si } r > 1. \end{cases}$$

• **Cas 2 :  $\xi < \xi_c$  (cellule bleue).**

En calculant un développement limité de  $|R|^2$  en 0, nous obtenons

$$|\bar{R}(i\Omega)| = 1 + (2\xi^2 - 1)(r - 1)(r + 1)\Omega^2 + O(\Omega^4),$$

avec  $2\xi^2 - 1 < 0$  et  $r + 1 > 0$ . Nous pouvons alors montrer que  $|\bar{R}(i\Omega)|$  est décroissante au voisinage de 0 si  $r > 1$ , et est croissante dans le cas contraire. En identifiant la valeur à l'infini de  $|\bar{R}(i\Omega)|$  suivant les valeurs de  $r$ , on identifie la pulsation du maximum de  $|\bar{R}(i\Omega)|$ , c'est-à-dire la pulsation à la résonance  $\Omega_r$ . On en déduit ainsi la norme  $H_\infty$  de  $\bar{R}$ . Les variations de  $|\bar{R}|$  sont résumées dans la Figure 5.5, ainsi que la valeur de la norme  $H_\infty$  de  $\bar{R}$  suivant les valeurs de  $r$ .

	$\Omega_0 = 0$	$\Omega_{-1}$	$\Omega_1$	$\infty$
$ \bar{R} $ si $r > 1$	1		$\ \bar{R}\ _\infty =  \bar{R}(\Omega_1)  > 1$	$r^2 > 1$
		$ \bar{R}(\Omega_{-1})  < 1$		
$ \bar{R} $ si $r < 1$	1	$\ \bar{R}\ _\infty =  \bar{R}(\Omega_{-1})  > 1$	$ \bar{R}(\Omega_1)  < 1$	$r^2 < 1$

FIGURE 5.5 – Variations de  $|\bar{R}(i\Omega)|$  pour  $\xi < \sqrt{2}/2$

Le remplacement des expressions symboliques des  $\Omega_\varepsilon$  dans  $|\bar{R}(i\Omega)|$  (et leur simplification via un logiciel de calcul symbolique) permet de conclure la preuve.

Pour se rendre mieux compte de la forme de  $\bar{R}$  dans ces deux cas, dans la Figure 5.6, nous avons tracé son diagramme de Bode dans chacun des cas.

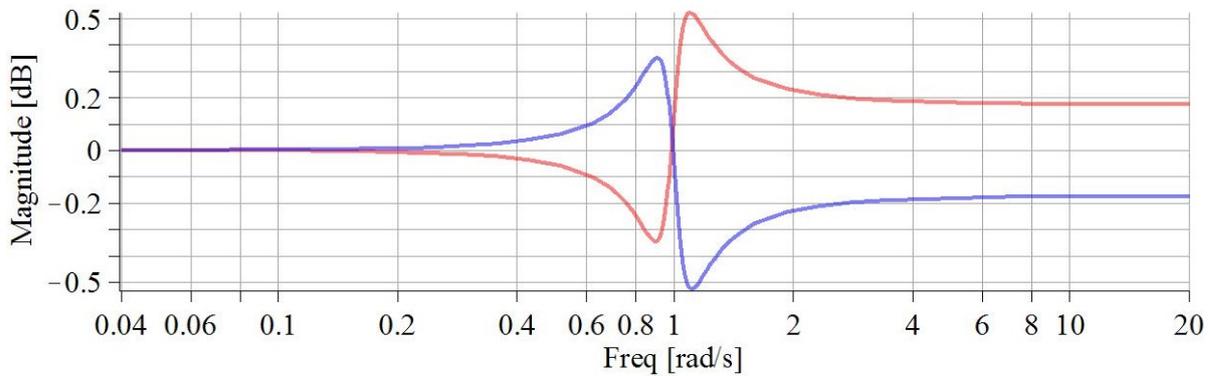


FIGURE 5.6 –  $|\bar{R}(i\Omega)|$  - Rouge :  $\xi = 0.1$ ,  $r = 1.01$  - Bleue :  $\xi = 0.1$ ,  $r = 0.99$

□

**Exemple 5.1.** Nous utilisons les valeurs numériques suivantes :

$$\omega_0 = 1, \quad \omega_1 = 1.03, \quad \xi = 1.08 \cdot 10^{-2}.$$

En évaluant le résultat symbolique obtenu en (5.23), nous obtenons  $\|R\|_\infty \simeq 3.1558$ . Une vérification de ce calcul sous MATLAB à l'aide de la commande `hinfnorm` donne  $\|R\|_\infty = 3.1546$ . On peut aussi tracer le diagramme de Bode de  $R$  et vérifier à nouveau ce calcul (voir Figure 5.7).

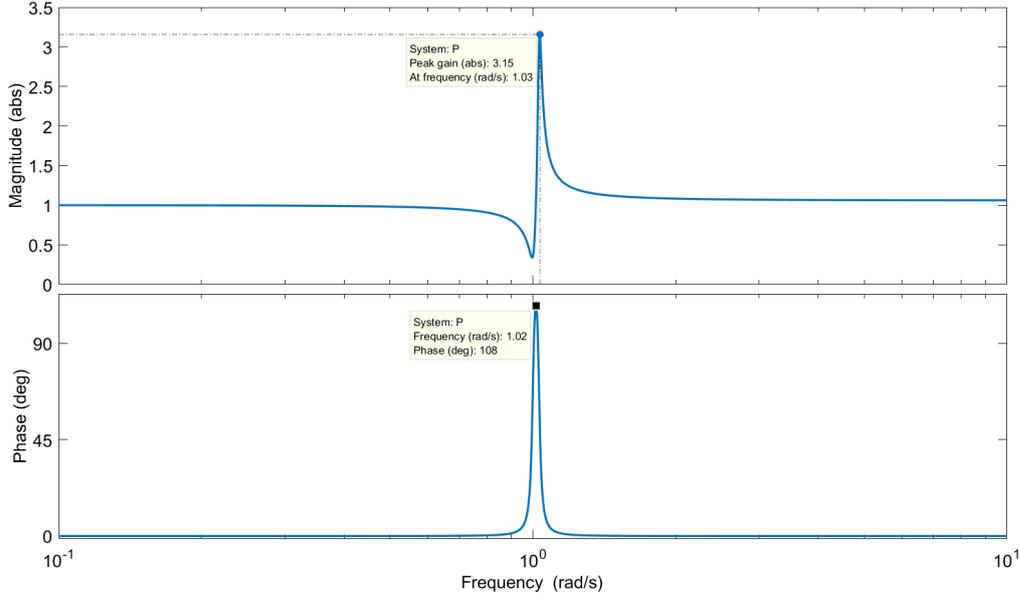


FIGURE 5.7 – Diagramme de Bode de  $R$  et sa norme  $H_\infty$

**Corollaire 5.1.** Soient  $r := \frac{\omega_1}{\omega_0}$ ,  $\xi_c := \frac{\sqrt{2}}{2}$  et  $\delta := (r^2 + 1)^2 - 4r^2(2\xi^2 - 1)^2$ . Le carré de la norme  $H_\infty$  de  $D$  est donnée par :

$$\|D\|_\infty^2 = \begin{cases} 1 + \frac{g^2 \tau^2}{2} + \sqrt{\left(1 + \frac{g^2 \tau^2}{2}\right)^2 - 1} & \text{si } \xi \geq \xi_c, r < 1, \\ 1 + \frac{g^2 \tau^2 r^4}{2} + \sqrt{\left(1 + \frac{g^2 \tau^2 r^4}{2}\right)^2 - 1} & \text{si } \xi \geq \xi_c, r > 1, \\ 1 + \frac{g^2 \tau^2 r^2}{2} \left(\frac{1 - r^2 + \sqrt{\delta}}{r^2 - 1 + \sqrt{\delta}}\right) + \sqrt{\left(1 + \frac{g^2 \tau^2 r^2}{2} \left(\frac{1 - r^2 + \sqrt{\delta}}{r^2 - 1 + \sqrt{\delta}}\right)\right)^2 - 1} & \text{si } \xi < \xi_c, r < 1, \\ 1 + \frac{g^2 \tau^2 r^2}{2} \left(\frac{r^2 - 1 + \sqrt{\delta}}{1 - r^2 + \sqrt{\delta}}\right) + \sqrt{\left(1 + \frac{g^2 \tau^2 r^2}{2} \left(\frac{r^2 - 1 + \sqrt{\delta}}{1 - r^2 + \sqrt{\delta}}\right)\right)^2 - 1} & \text{si } \xi < \xi_c, r > 1. \end{cases}$$

*Démonstration.* Ce corollaire est une conséquence directe du Théorème 5.3, du Lemme 5.1 et de la Proposition 5.2.  $\square$

Nous obtenons alors un encadrement de  $\|\Pi(K, G)\|_\infty$  :

$$\frac{\|\Pi(K_0, G_0)\|_\infty}{\|D\|_\infty^2} \leq \|\Pi(K, G)\|_\infty \leq \|D\|_\infty^2 \|\Pi(K_0, G_0)\|_\infty.$$

Disposant d'une forme close de  $\gamma_{0_{\text{opt}}}$  tel que

$$\|\Pi(K_0, G_0)\|_{\infty} \leq \gamma_{0_{\text{opt}}},$$

nous obtenons alors une formule explicite d'une borne supérieure de  $\|\Pi(K, G)\|_{\infty}$  d'un système d'ordre 4 avec retard, c'est-à-dire :

$$\|\Pi(K, G)\|_{\infty} \leq \|D\|_{\infty}^2 \gamma_{0_{\text{opt}}}.$$

**Exemple 5.2.** Nous utilisons les valeurs numériques suivantes :

$$g_0 = 6.65 \cdot 10^{-3}, \quad \tau = 1.21, \quad \omega_0 = 1, \quad \omega_1 = 1.03, \quad \xi = 1.08 \cdot 10^{-2}.$$

Nous obtenons :

$$0.976 \|\Pi(K_0, G_0)\|_{\infty} \leq \|\Pi(K, G)\|_{\infty} \leq 1.025 \|\Pi(K_0, G_0)\|_{\infty}.$$

La borne obtenue sur  $\|\Pi(K, G)\|_{\infty}$  reste proche de  $\|\Pi(K_0, G_0)\|_{\infty}$ . Ceci est dû au faible gain  $g_0$ . Ainsi, la robustesse de l'asservissement du système  $G$  bouclé par  $K$  est identique à celle du système  $G_0$  bouclé par  $K_0$ .

Par ailleurs pour un tel système, en utilisant les formules explicites de  $\gamma_{\text{opt}}$  de systèmes d'ordre 1 à 4 obtenues dans la Section 4.1.1, nous obtenons :

$$\|\Pi(K_0, G_0)\|_{\infty} \simeq 1.42.$$

Encore,  $\|\Pi(K_0, G_0)\|_{\infty}$  est très faible à cause du gain faible  $g_0$ . Nous obtenons alors :

$$\|\Pi(K, G)\|_{\infty} \leq 1.46.$$

Cette borne sur  $\|\Pi(K, G)\|_{\infty}$  montre que le système en boucle fermée obtenu est très robuste voir trop robuste, en d'autres termes possède de très grandes marges de stabilité. Ceci est dû au fait que  $g_0$  est très faible. En pratique, une telle borne ne permettrait pas d'obtenir un asservissement de  $G$  suffisamment performant car trop robuste. Il faut pour cela combiner cette approche avec la méthode de Loop-Shaping.

La section suivante aborde la problématique de Loop-Shaping dans le contexte des systèmes à retard décomposés additivement.

## 5.3 Réglage d'un correcteur robuste sous-optimal à l'aide d'une pondération Loop-Shaping

Dans cette section, nous proposons de combiner l'inversion de pondération développée dans la Section 4.3 avec les autres résultats de ce chapitre sur les systèmes à retard et la décomposition additive. L'objectif est d'obtenir de meilleures performances en boucle fermée par réglage d'une pondération scalaire.

Comme nous allons proposer de régler une pondération Loop-Shaping, nous allons bien définir nos systèmes et distinguer le système réel et le système fictif. Nous verrons ensuite comment exploiter les connaissances que l'on a sur le système afin de régler cette pondération de manière à obtenir un système bouclé avec retard conservant des propriétés de robustesse lorsqu'un des paramètres varie.

### 5.3.1 Système réel et hypothèses

Nous proposons de travailler à nouveau sur le système à retard  $G_{4,r}$ . Nous l'appelons *système réel* et nous le notons par  $P$ . Nous rappelons ici sa forme

$$P := P_0 e^{-\tau s},$$

où :

$$P_0 = \frac{p}{s} \left( \frac{\frac{s^2}{\omega_0^2} + 2\xi \frac{s}{\omega_0} + 1}{\frac{s^2}{\omega_1^2} + 2\xi \frac{s}{\omega_1} + 1} \right) \frac{1}{Ts + 1}.$$

Nous supposons que les paramètres suivants sont fixés :

$$\omega_0 := 1, \quad T := 1/1.14, \quad \tau := 1.21, \quad \xi := 1.08 \cdot 10^{-2}, \quad p := 6.65 \cdot 10^{-3}. \quad (5.25)$$

Nous supposons aussi que la fréquence modale naturelle  $\omega_1$  peut varier très lentement suivant les paramètres et est considérée comme un paramètre non fixé. De plus, nous supposons que  $\omega_1$  satisfait l'inégalité suivante :

$$\omega_1 > 1.$$

La Figure 5.8 représente le diagramme de Bode de  $P$  lorsque la fréquence  $\omega_1$  prend des valeurs dans l'ensemble suivant :

$$\Omega_* := \left\{ \begin{array}{l} 1.1, 1.53, 2.00, 2.47, 2.94, 3.41, 3.88, 4.35, 4.82, 5.29, \dots \\ \dots, 5.76, 6.24, 6.71, 7.18, 7.65, 8.12; 8.59, 9.06, 9.53, 10 \end{array} \right\}.$$

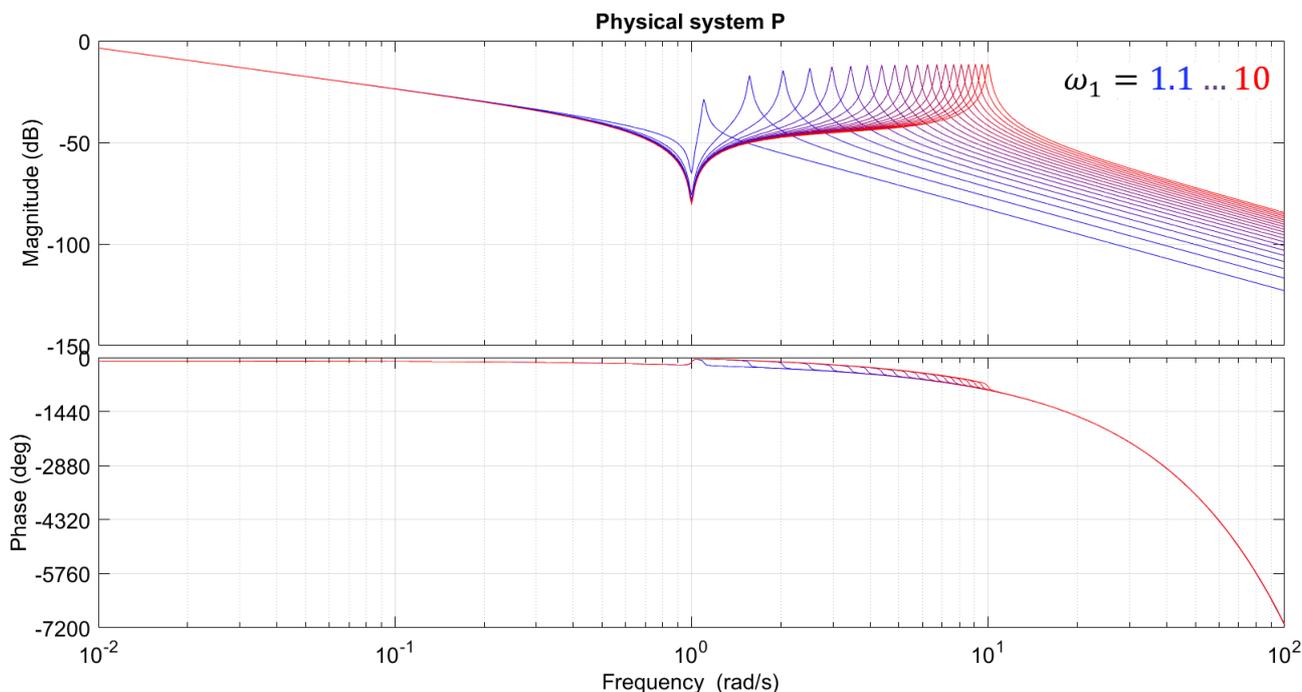


FIGURE 5.8 – Diagramme de Bode du système  $P$  pour  $\omega_1 \in \Omega_*$

### 5.3.2 Système fictif et données préliminaires

Comme dans le cadre du Loop-Shaping étudié dans ce mémoire, au système réel  $P$ , nous associons le *système fictif*  $G$  défini par :

$$G := w P.$$

Nous définissons le gain statique de  $G$  par :

$$g = w p.$$

Comme  $G = w P_0 e^{-\tau s}$ , nous définissons aussi

$$G_0 := w P_0,$$

et nous avons alors :

$$G := G_0 e^{-\tau s}.$$

Voir Figure 5.14 pour une vision globale des transformations.

**Remarque 5.5.** Avec ces notations,  $G$  (respectivement  $G_0$ ) est obtenu en substituant  $p$  par  $g$  dans  $P$  (respectivement  $P_0$ ).

Notons par  $\theta$  les paramètres non fixés du système fictif :

$$\theta := (g, \omega_1).$$

Comme  $G_0$  est un système SISO d'ordre 4, nous pouvons trouver un correcteur  $K_0$  stabilisant  $G_0$  dépendant explicitement des paramètres  $\theta$  (voir Section 4.1.2). De plus, nous pouvons aussi obtenir une borne supérieure  $\gamma_{0,\text{opt}}(\theta)$  sur  $\|\Pi(K_0, G_0)\|_\infty$  telle que :

$$\gamma_{0,\text{opt}}(\theta) := \min_{K_0(\theta)} \{ \gamma_0 \mid \|\Pi(K_0(\theta), G_0)\|_\infty \leq \gamma_0 \}.$$

Cette borne dépend explicitement de  $\theta$ . Nous avons tracé celle-ci dans la Figure 5.9. Par conséquent, les marges de stabilité  $m_G$  et  $m_\Phi$  de  $K_0 G_0$  assurées par  $\gamma_{0,\text{opt}}(\theta)$  sont elles aussi connues explicitement (voir (2.10) pour un rappel sur la définition de ces dernières).

D'après le Théorème 5.1, si  $K_0$  stabilise  $G_0$ , alors  $K := \frac{K_0}{1+K_0\Delta}$  stabilise le système fictif  $G$ , ce qui implique que le correcteur  $C = w K$  stabilise  $P$  (voir Section 2.5 et Figure 5.14).

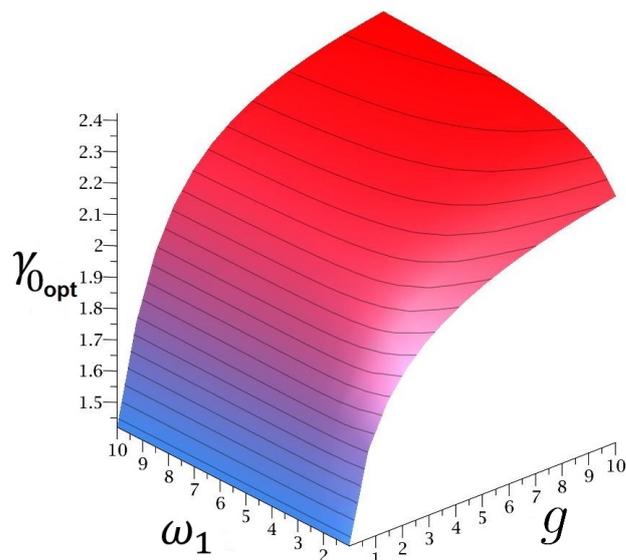
Nous cherchons maintenant une méthode pour régler la pondération  $w$  afin que le critère

$$\gamma_{\text{opt}}(\theta) := \min_{K(\theta)} \{ \gamma \mid \|\Pi(K(\theta), G)\|_\infty \leq \gamma \}$$

atteigne une valeur désirée  $\bar{\gamma}$ . Le problème majeur est que nous n'avons aucune expression explicite de  $\gamma_{\text{opt}}(\theta)$ , ni d'une borne exploitable sur celui-ci. En effet, pour  $w = 564.7$  et  $\omega_1 = 2$ , la borne (5.22) donne :

$$\|\Pi(K, G)\|_\infty \leq 538 \|\Pi(K_0, G_0)\|_\infty.$$

Bien que le diagramme de la boucle ouverte (voir Figure 5.10) confirme que les marges de stabilité obtenues pour le système à retard  $PC$  soient acceptables, la borne (5.22) ne nous permet pas de confirmer ces marges de stabilité car celle-ci est bien trop large. En d'autres


 FIGURE 5.9 – Tracé de  $\gamma_{0_{\text{opt}}}(\theta)$ 

termes, nous avons un moyen de concevoir un correcteur dépendant explicitement des paramètres du système et stabilisant le système avec retard mais nous ne savons pas si celui-ci conserve certaines propriétés de robustesse, héritées du système bouclé sans retard, lorsque les paramètres varient.

La sous-section suivante permet d'observer que, lorsque l'on règle une pondération  $w$  de telle sorte que  $\gamma_{0_{\text{opt}}}(\theta)$  reste proche d'une valeur souhaitée, alors le système avec retard conserve des propriétés de robustesse. De plus, la valeur de  $\gamma_{0_{\text{opt}}}(\theta)$  constitue un excellent moyen de réglage afin de régler la robustesse du système bouclé avec retard. Ce résultat a été observé par des essais empiriques dont la justification reste à prouver. Cette méthode reste donc à développer théoriquement. Exposons-en le principe en détail et les caractéristiques observées.

### 5.3.3 Une procédure de conception de correcteur Loop-Shaping adaptatif stabilisant le $G$

Dans un premier temps, supposons  $\omega_1$  fixé à la valeur suivante

$$\bar{\omega}_1 := 2,$$

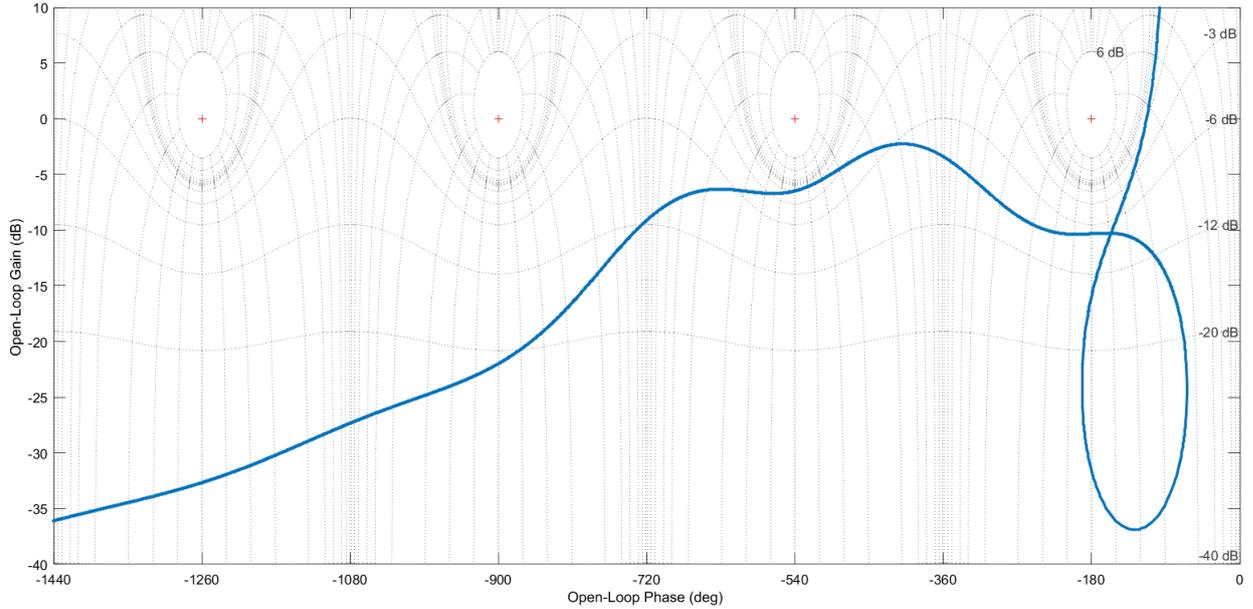
les autres paramètres physiques étant eux toujours fixés aux valeurs données par (5.25). La pondération  $w$  est toujours non fixée. En utilisant un diagramme de Black-Nichols de la boucle ouverte du système avec retard  $PC$ , nous réglons par essai-erreur la pondération  $w$  de telle sorte que les marges de stabilité  $M_G$  et  $M_\phi$  de la boucle ouverte  $PC$  soient suffisamment grandes. La valeur obtenue pour un tel  $w$  est

$$w = \bar{w} = 82.82,$$

ce qui donne les marges de stabilité suivantes :

$$\begin{cases} M_G(P, C) = 10.3 \text{ dB}, \\ M_\Phi(P, C) = 66.7^\circ. \end{cases}$$

Ces marges sont visibles sur la Figure 5.10.


 FIGURE 5.10 – Tracé du diagramme de Black-Nichols de  $PC$  pour  $w = 82.82$  et  $\omega_1 = 2$ 

Pour cette valeur de pondération  $w$ , la valeur correspondante de  $\gamma_0^{\text{opt}}$  est 1.7, et nous avons :

$$\|\Pi(K_0, G_0)\|_\infty \leq 1.7.$$

Supposons maintenant que  $\omega_1$  est non fixé à une valeur numérique mais vérifie la condition  $\omega_1 > 1$ . Soit  $\bar{\gamma}_0$  une valeur de  $\gamma_{0,\text{opt}}$  choisie. La remarque précédente a suggéré que

$$\bar{\gamma}_0 = 1.7,$$

permettait d'atteindre des bonnes marges de stabilité pour la boucle ouverte du système avec  $PC$ . Nous conjecturons que, si l'on maintient  $\gamma_{0,\text{opt}}$  proche de  $\bar{\gamma}_0$  lorsque les paramètres varient, alors les marges de stabilité de  $PC$  varient faiblement. Dans un premier temps, nous cherchons donc à régler  $w$  de telle sorte que  $\gamma_{0,\text{opt}}$  reste proche de  $\bar{\gamma}_0$  pour tout  $\omega_1 > 1$ . Ce travail a déjà été fait dans la Section 4.3 pour d'autres systèmes. Concrètement, en intersectant la surface représentée en Figure 5.9 par le plan  $\gamma_{0,\text{opt}} = \bar{\gamma}_0$ , nous obtenons la courbe bleue de la Figure 5.11. C'est cette même courbe que nous cherchons à suivre pour tout  $\omega_1$ . Obtenir une expression explicite de la forme  $g = g_w(\omega_1)$  permettant d'obtenir la courbe bleue n'est pas possible dans ce cas (le polynôme  $L$  défini en (4.25) est de degré 68 en  $g$ ). Nous cherchons alors à approximer localement cette courbe par une fonction de la forme :

$$g = \hat{g}(\omega_1).$$

L'approximation suivante, obtenue à partir de développements en série de Taylor pour le polynôme  $L(\lambda, \frac{1}{\omega_1}, g)$  (défini en (4.25)) particularisé en  $\lambda = \bar{\gamma}_0^2 - 1$ , permet d'approximer cette courbe autour du point  $(\omega_1, g) = (\bar{\omega}_1, \bar{w}p) = (2, 0.5647)$  :

$$\hat{g}(\omega_1) := 0.7170 - \frac{0.9320}{\omega_1} + \frac{1.5150}{\omega_1^2} - \frac{0.6320}{\omega_1^3}.$$

La courbe  $g = \hat{g}(\omega_1)$  est représentée en rouge dans la Figure 5.11.

Nous pouvons vérifier que le choix de  $g = \widehat{g}(\omega_1)$  permet d'obtenir  $\gamma_{0,\text{opt}}(g, \omega_1)$  proche de  $\bar{\gamma}_0$  autour du point de fonctionnement choisi. Il suffit pour cela d'évaluer la fonction explicite  $\gamma_{0,\text{opt}}(\widehat{g}(\omega_1), \omega_1)$  fonction uniquement de la variable  $\omega_1$  (Figure 5.12). En particulier  $\gamma_{0,\text{opt}}$  vaut exactement  $\bar{\gamma}_0$  au point  $\omega_1 = \bar{\omega}_1$  choisi et reste proche de la valeur choisie lorsque l'on s'éloigne de ce point. Plus l'approximation  $g_{\bar{\gamma}_0}(\omega_1)$  est fine, plus  $\gamma_{0,\text{opt}}$  reste proche de la valeur souhaitée.

En choisissant  $\gamma > \gamma_{0,\text{opt}}(\widehat{g}(\omega_1), \omega_1)$ , nous construisons alors un correcteur  $H_\infty$  adaptatif  $K_0(s, \omega_1)$  stabilisant  $G_0(s, \omega_1)$ , dépendant explicitement du paramètre  $\omega_1$  non fixé et qui assure que  $\gamma_{0,\text{opt}}$  reste proche de  $\bar{\gamma}_0$  autour de  $\omega_1 = \bar{\omega}_1$ . En utilisant le Théorème 5.1, le correcteur suivant

$$K(s, \omega_1) = \frac{K_0(s, \omega_1)}{1 + K_0(s, \omega_1) \Delta(s, \omega_1)}$$

stabilise  $G(s, \omega_1)$ . En choisissant la pondération

$$w = \widehat{w}(\omega_1) := \frac{\widehat{g}}{g},$$

nous en déduisons alors un correcteur

$$C(s, \omega_1) := \widehat{w}(\omega_1) K(s, \omega_1) = \frac{\widehat{w}(\omega_1) K_0(s, \omega_1)}{1 + K_0(s, \omega_1) \Delta(s, \omega_1)}$$

stabilisant  $P(s, \omega_1)$ . Les valeurs prises par la pondération scalaire  $\widehat{w}(\omega_1)$  sont tracées dans la Figure 5.13.

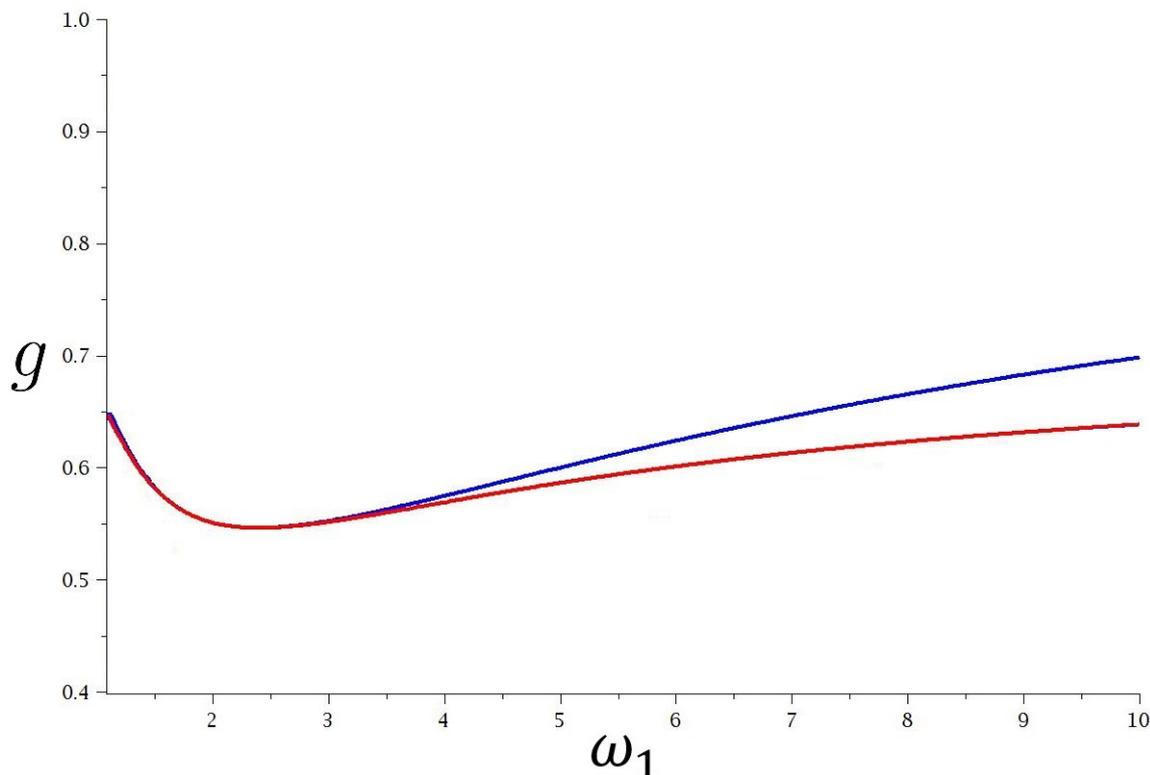


FIGURE 5.11 – Bleue :  $\gamma_{0,\text{opt}}(g, \omega_1) = 1.7$  – Rouge :  $g = \widehat{g}(\omega_1)$

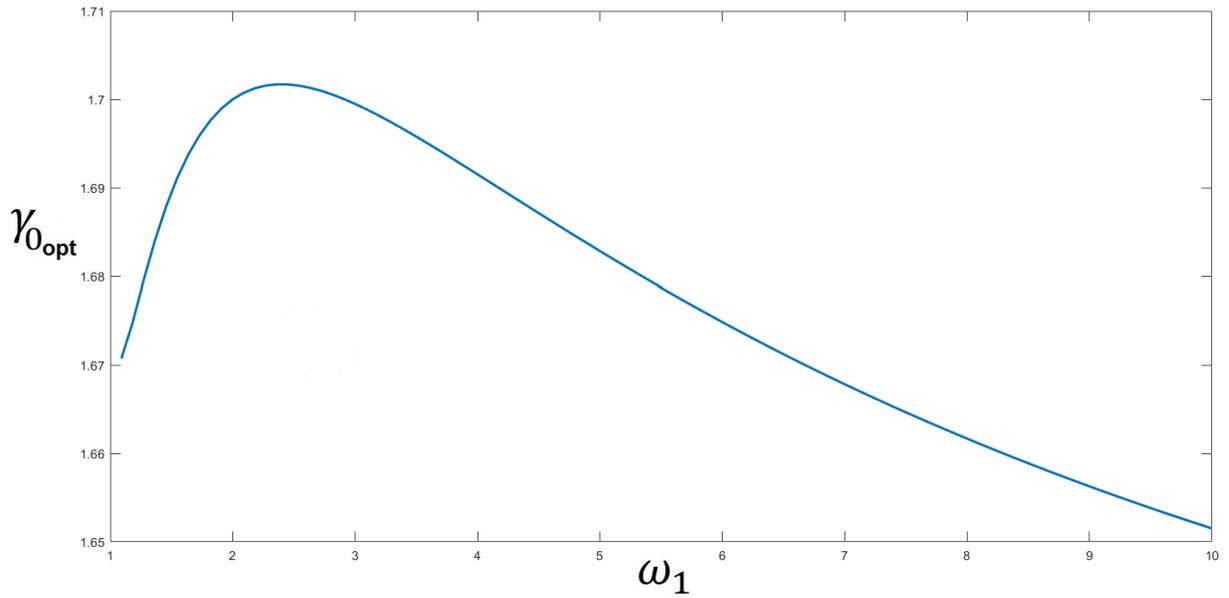


FIGURE 5.12 – Graphe de  $\gamma_{0_{\text{opt}}}(\hat{g}(\omega_1), \omega_1)$

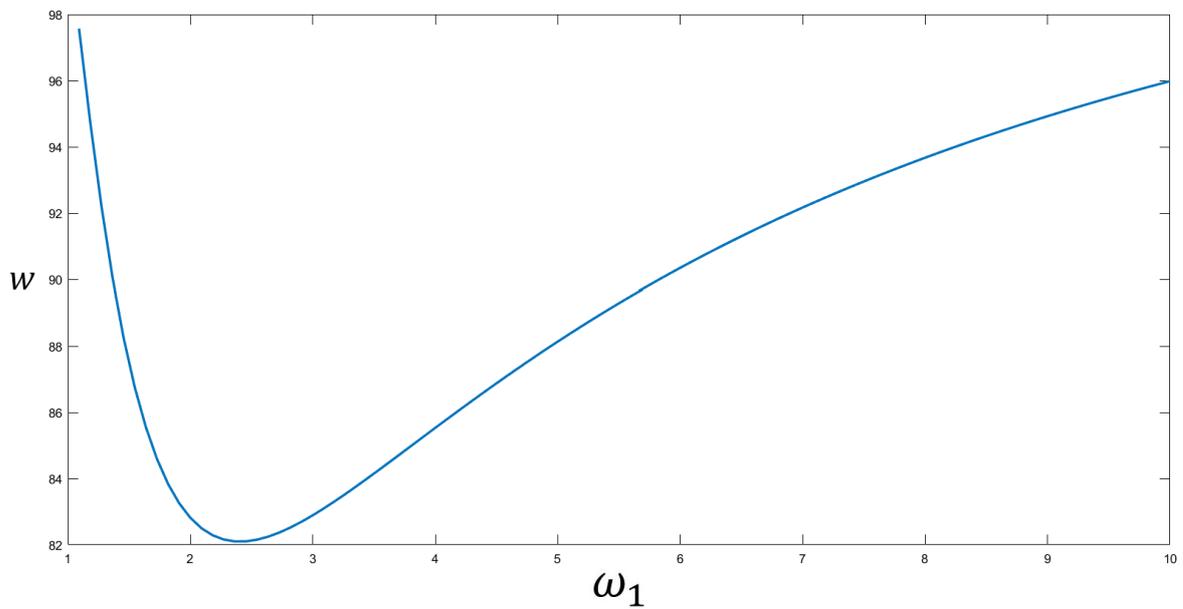


FIGURE 5.13 – Graphe de  $\hat{w}(\omega_1)$

La procédure de conception de correcteur par Loop-Shaping est résumée par le schéma de la Figure 5.14. Sur ce schéma, le tracé en trait plein correspond à la procédure décrite ci-dessus tandis que le trait discontinu représente une procédure duale permettant d'arriver au même résultat. Récapitulons les étapes de cette méthode en s'appuyant du schéma de la Figure 5.14 :

1.  $P$  est le système réel avec retard à asservir. De ce système, nous définissons un système fictif  $G$  à l'aide d'une pondération  $w$ .
2. Nous décomposons le système fictif avec retard  $G$  en une partie sans retard  $G_0$  et une partie stable  $\Delta$  contenant le retard distribué.
3. Le système sans retard  $G_0$  étant d'ordre 4, nous disposons alors d'une expression explicite de  $\gamma_{0\text{opt}}$  en fonction des paramètres du modèle. Choisisant  $\gamma > \gamma_{0\text{opt}}$ , nous en déduisons un correcteur  $K_0$  stabilisant  $G_0$ . A l'aide d'une étude préalable de la courbe  $\gamma_{0\text{opt}} = \bar{\gamma}_0$ , nous choisissons  $w$  permettant de conserver  $\gamma_{0\text{opt}}$  proche de  $\bar{\gamma}_0$  autour d'un point de fonctionnement choisi.  $\gamma_{0\text{opt}}$  **constitue un moyen de réglage, réglé via  $w$ , permettant de contrôler la robustesse du système bouclé avec retard.**
4. En appliquant le Théorème 5.1, nous en déduisons un correcteur  $K$  stabilisant le système fictif avec retard  $G$ .
5. En basculant la pondération dans le correcteur, nous déduisons alors un correcteur  $C$  stabilisant le système réel avec retard  $P$ .

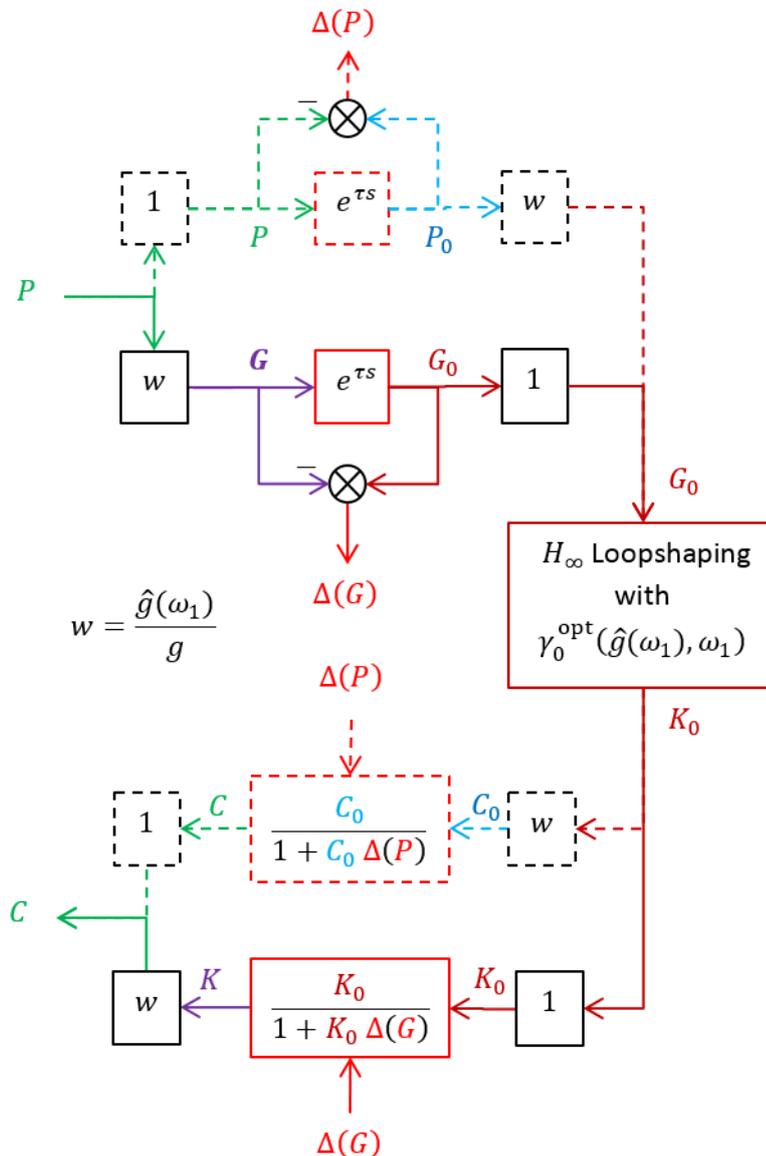


FIGURE 5.14 – Procédure de Loop-Shaping pour un système à retard

La Figure 5.15 représente le diagramme de Bode du correcteur  $C(s, \omega_1)$  obtenu pour différentes valeurs de  $\omega_1$ . En vert, nous avons représenté le correcteur nominal, c'est-à-dire celui obtenu pour  $\omega_1 = \bar{\omega}_1$ . Nous remarquons que  $C$  le plus faible gain à la fréquence du mode et que cette baisse de gain s'adapte suivant la fréquence du mode. C'est ce que l'on appelle un *réjecteur*. Un gros avantage du correcteur explicite  $C$  obtenu est que celui-ci s'adapte automatiquement aux variations du système sans avoir besoin d'être recalculé par un algorithme d'optimisation.

### 5.3.4 Caractéristiques de l'asservissement obtenu

Le diagramme de Bode de la boucle ouverte obtenue est donné en Figure 5.16 et son diagramme de Black-Nichols en Figure 5.17.

Commençons par commenter la Figure 5.16. Nous remarquons que la bande-passante est grande pour  $\omega_1$  très proche de  $\omega_0$  et qu'elle reste très proche de 0.3 lorsque  $\omega_1$  s'écarte de cette valeur.

Regardons maintenant la Figure 5.17. Cette figure est la plus intéressante de notre point de vue. Sur celle-ci, nous pouvons observer que les marges de stabilité de la boucle ouverte sont conservées lorsque  $\omega_1$  varie. Plus spécifiquement, pour les points choisis pour le tracé tels que  $1.1 < \omega_1 < 10$ , nous avons :

$$\begin{aligned} M_G(C, P) &> 9 \text{ dB}, \\ M_\Phi(C, P) &> 70^\circ. \end{aligned}$$

Ceci est très intéressant du point de vue pratique. En effet, pour cet exemple, nous avons considéré de très larges variations paramétriques et la stabilité du système bouclé semblent ne pas changer lorsque l'on fait varier  $\omega_1$ .

A titre de comparaison, considérons le correcteur nominal obtenu pour  $\omega_1 = \bar{\omega}_1 = 2$  (voir la courbe verte de la Figure 5.15). En faisant varier les paramètres de  $P$  mais sans changer de correcteur, c'est-à-dire en gardant le correcteur nominal, traçons les variations de la boucle ouverte  $CP$  sur un diagramme de Black-Nichols dans la Figure 5.18. Nous remarquons que le système bouclé devient très vite instable du fait des variations de  $\omega_1$  ! C'est cette approche qui est employée en pratique au sens où l'ingénieur conçoit un correcteur pour un système nominal en faisant en sorte que celui-ci stabilise une boule de modèles particulière. Il doit cependant réitérer la conception de son correcteur à chaque fois que la boule de modèles à stabiliser change, et/ou alors considérer une grosse boule de modèles. Cette dernière option a le désavantage d'obtenir un correcteur très robuste et donc peu performant. De plus, le calcul d'un nouveau correcteur à chaque changement de la boule de modèle est un processus itératif long. C'est cet aspect pratique qui a, en partie, motivé cette thèse.

Nous proposons maintenant de discuter des caractéristiques de la boucle fermée obtenue. Pour cela, nous revenons aux notions expliquées dans la Section 2.2 détaillant le rôle de chacune des fonctions de sensibilité  $S := \frac{1}{1+PC}$ ,  $PS$ ,  $CS$  et  $T := PCS$ . Celles-ci sont tracées dans les Figures 5.19–5.22. De plus, nous avons tracé les réponses à des échelons de  $PS$  et  $T$  en Figure 5.23 et Figure 5.24 respectivement. Nous remarquons plusieurs faits importants. Premièrement, nous avons  $20 \log(\|T\|_\infty) \leq 0 \text{ dB}$  pour tout  $\omega_1 \in \Omega_*$ , ce qui assure une bonne

robustesse à des incertitudes de modèle caractérisées par des erreurs multiplicatives (voir Section 2.2). Ensuite, nous avons  $20 \log(\|S\|_\infty) \leq 6$  dB pour tout  $\omega_1 \in \Omega_*$ , ce qui assure un bon suivi de consigne. Nous remarquons aussi de grands gains dans  $PS$  dus au fait que la structure mécanique comporte des modes résonnants. Ceux-ci ne peuvent être atténués par le correcteur  $C$  sans détériorer la fonction  $S$ . La commande  $H_\infty$  réalise donc un compromis “automatique”. Alors, nous constatons nécessairement des oscillations dans la réponse de  $PS$  à un échelon (Figure 5.23). En pratique, nous ajoutons des effets intégrateurs en basse fréquence dans  $C$  afin de faire tendre ces réponses vers zéro (rejet de perturbation). Ceci n’a pas été fait ici.

Il faut cependant “raisonner” l’approche que nous défendons ici. Lorsque nous traçons les Figures 5.16 et Figures 5.17, nous supposons que le système est parfaitement connu, c’est-à-dire que nous connaissons parfaitement  $\omega_1$  pour ensuite l’utiliser pour calculer le correcteur  $C$ . En pratique, ceci n’est jamais le cas, et cela justifie l’approche de la stabilisation de boules de modèle par critère  $H_\infty$ . En effet, dans cette approche classique, il suffit de considérer un modèle nominal et de concevoir un unique correcteur de ce système nominal. Il n’y a alors qu’à estimer d’éventuelles erreurs de modèle. Dans le cas que nous défendons ici, nous choisissons d’être plus proches de notre système afin d’adapter le correcteur à celui-ci. En pratique, il faut combiner nos méthodes avec une identification des paramètres du système afin de pouvoir estimer  $\omega_1$  et recalculer  $C$  automatiquement. Il existe pour cela une vaste littérature que nous avons commencé à étudier mais qui demanderait un travail plus approfondi pour établir un état de l’art pour notre problème. Par manque de temps, cet aspect n’a pas pu être approfondi dans cette thèse. Citons par exemple [LZ05] ou [GWY08] donnant plusieurs méthodes en boucle ouverte ou boucle fermée. Cependant, à notre connaissance, il n’existe aucune méthode à l’heure actuelle pour estimer les paramètres d’un système à retard, comportant des intégrateurs, en boucle fermée.

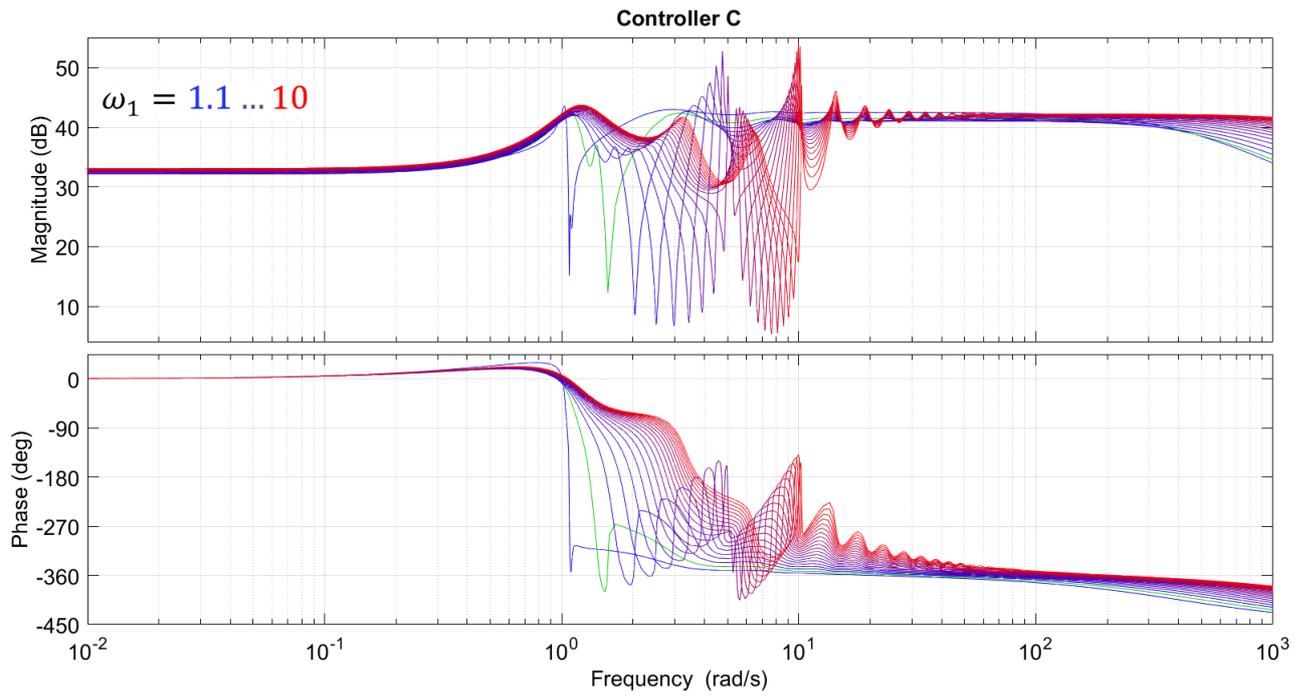


FIGURE 5.15 – Correcteur  $C(s, \omega_1)$  obtenu pour différentes valeurs de  $\omega_1$

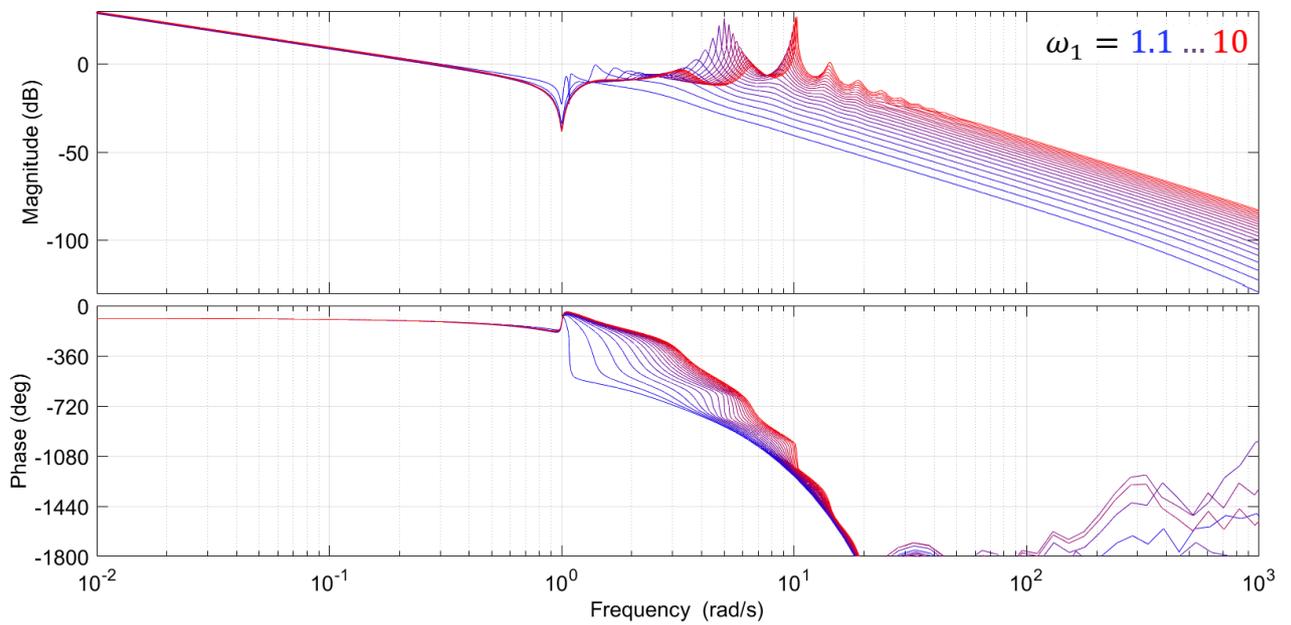


FIGURE 5.16 – Diagramme de Bode de la boucle ouverte  $C(\omega_1)P(\omega_1)$  obtenu pour  $\omega_1 \in \Omega_*$

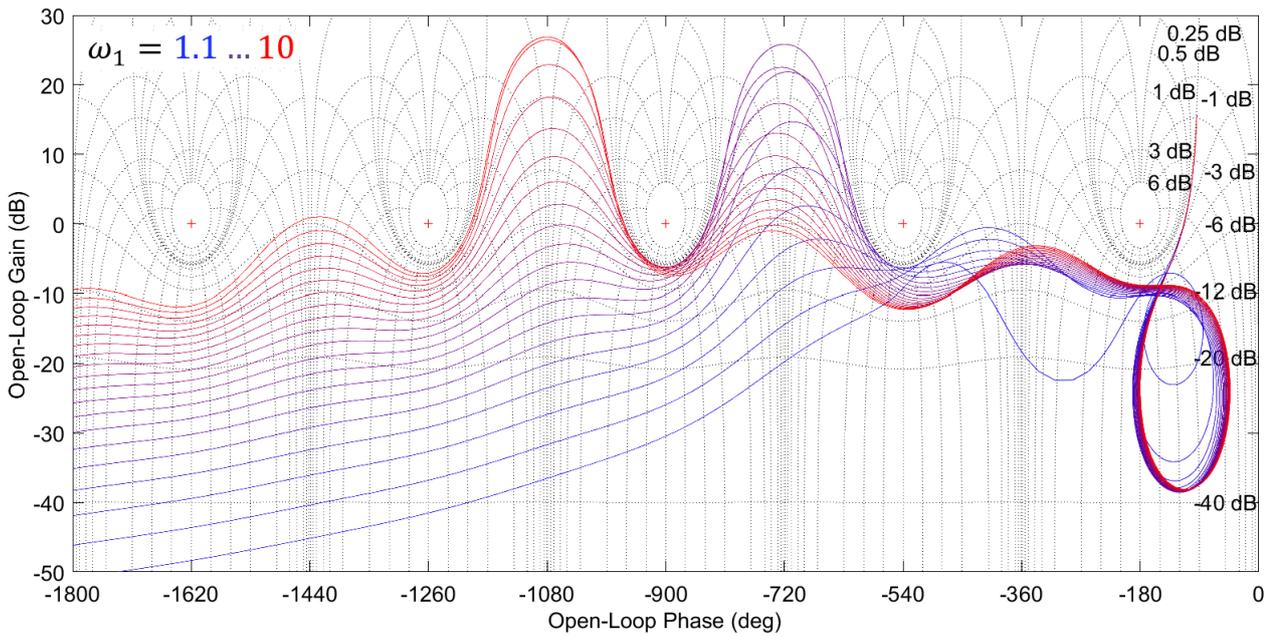


FIGURE 5.17 – Diagramme de Black-Nichols de  $C(\omega_1)P(\omega_1)$  obtenu pour  $\omega_1 \in \Omega_*$

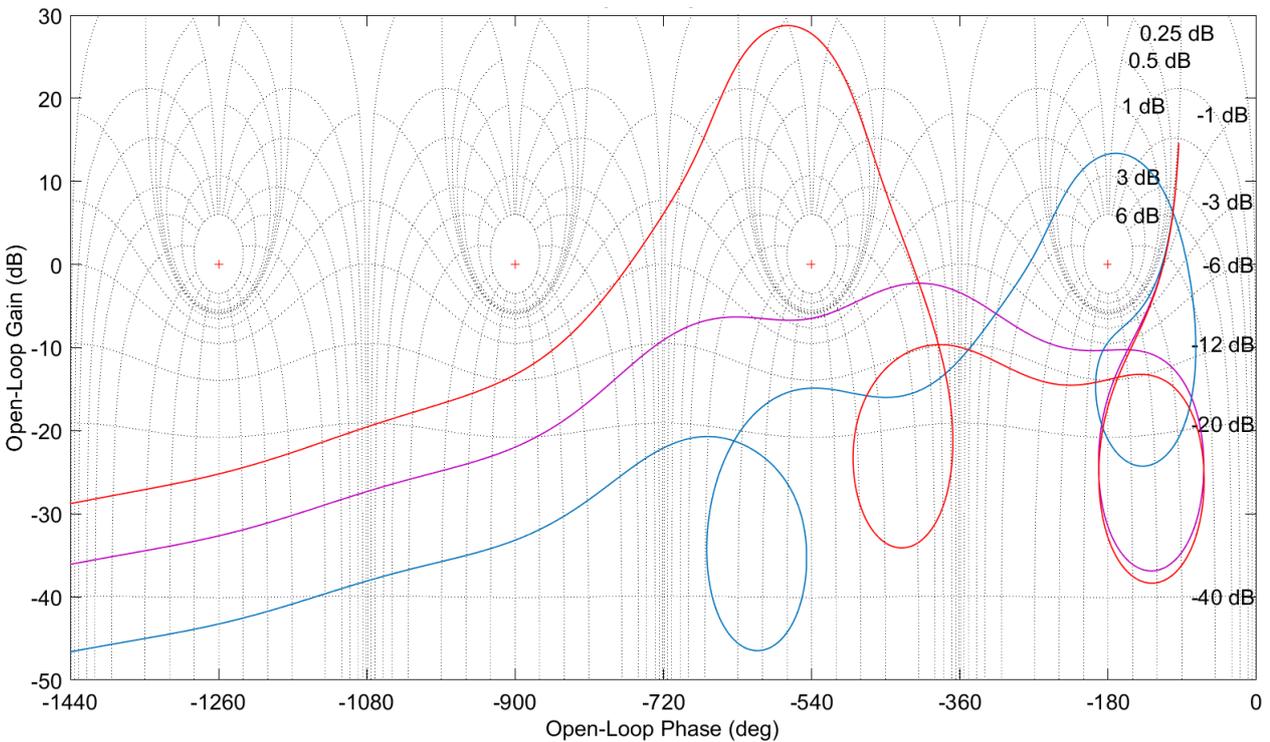


FIGURE 5.18 – Diagramme de Black-Nichols de  $C(\omega_1)P(\omega_1)$  obtenu pour différentes valeurs de  $\omega_1$  lorsque  $C$  est fixé au correcteur nominal. Bleu :  $\omega_1 = 1.1$  (instable) – Violet :  $\omega_1 = 2$  (stable) – Rouge :  $\omega_1 = 3$  (instable)

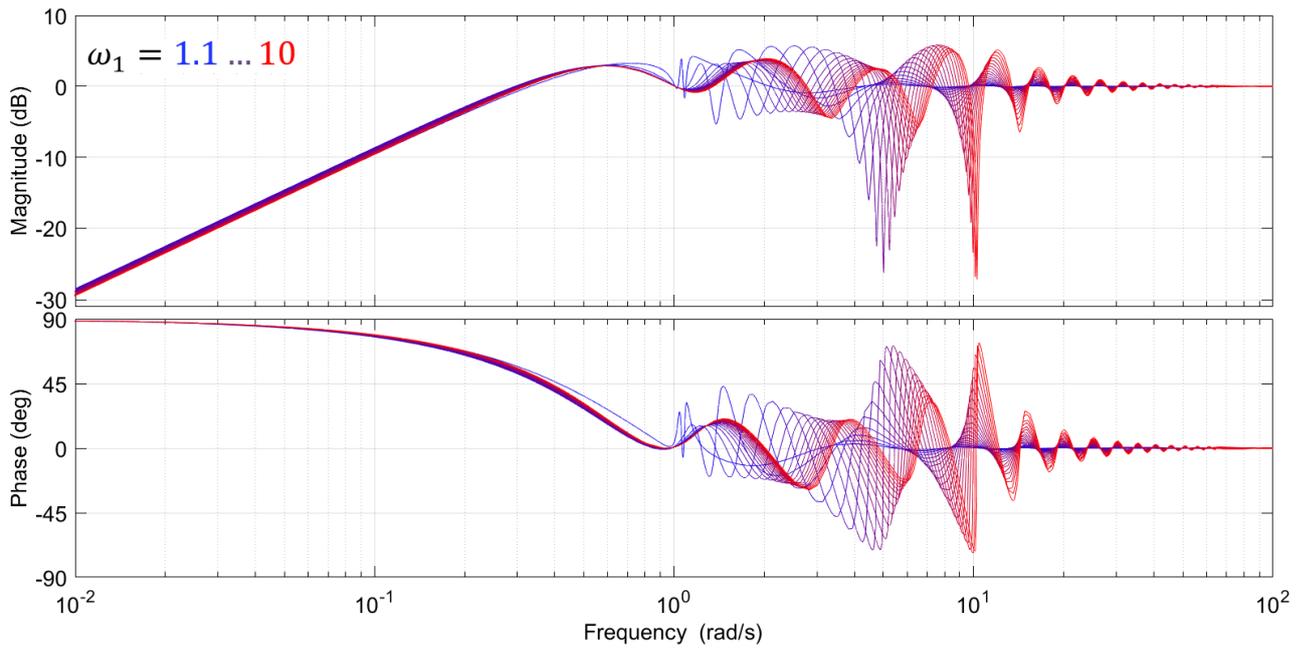


FIGURE 5.19 – Diagramme de Bode de  $S(\omega_1) = \frac{1}{1+C(\omega_1)P(\omega_1)}$  obtenu pour  $\omega_1 \in \Omega_*$

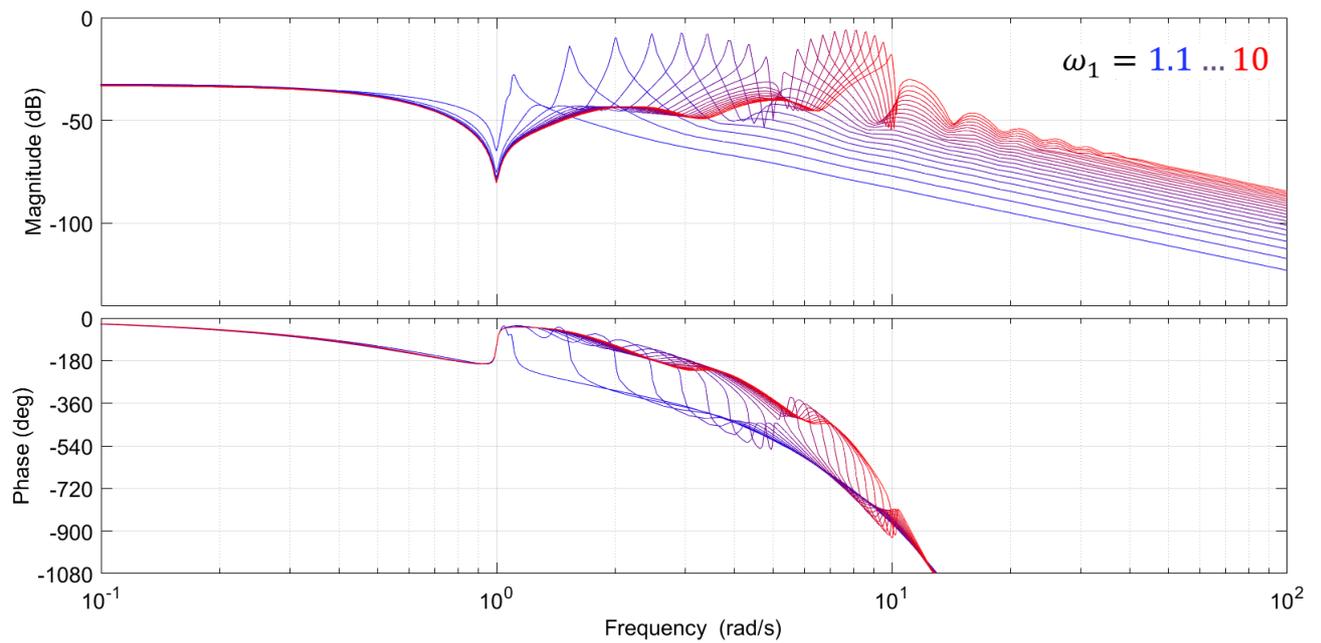


FIGURE 5.20 – Diagramme de Bode de  $P(\omega_1)S(\omega_1) = \frac{P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$  obtenu pour  $\omega_1 \in \Omega_*$

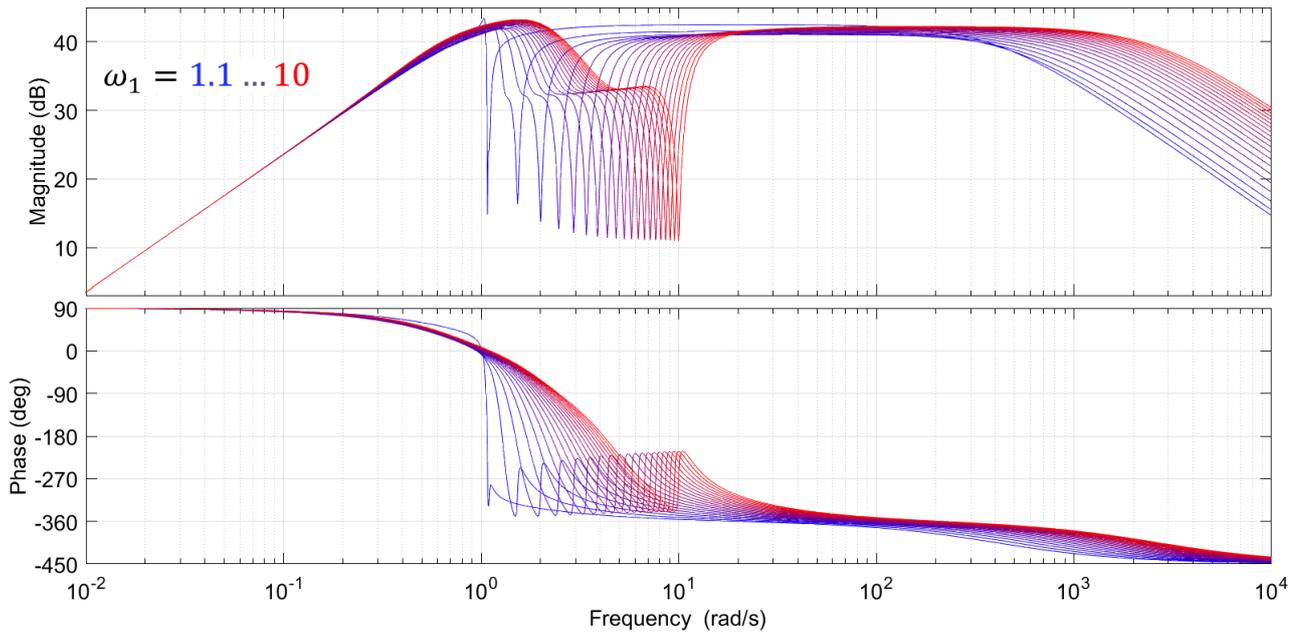


FIGURE 5.21 – Diagramme de Bode de  $C(\omega_1)S(\omega_1) = \frac{C(\omega_1)}{1+C(\omega_1)P(\omega_1)}$  obtenu pour  $\omega_1 \in \Omega_\star$

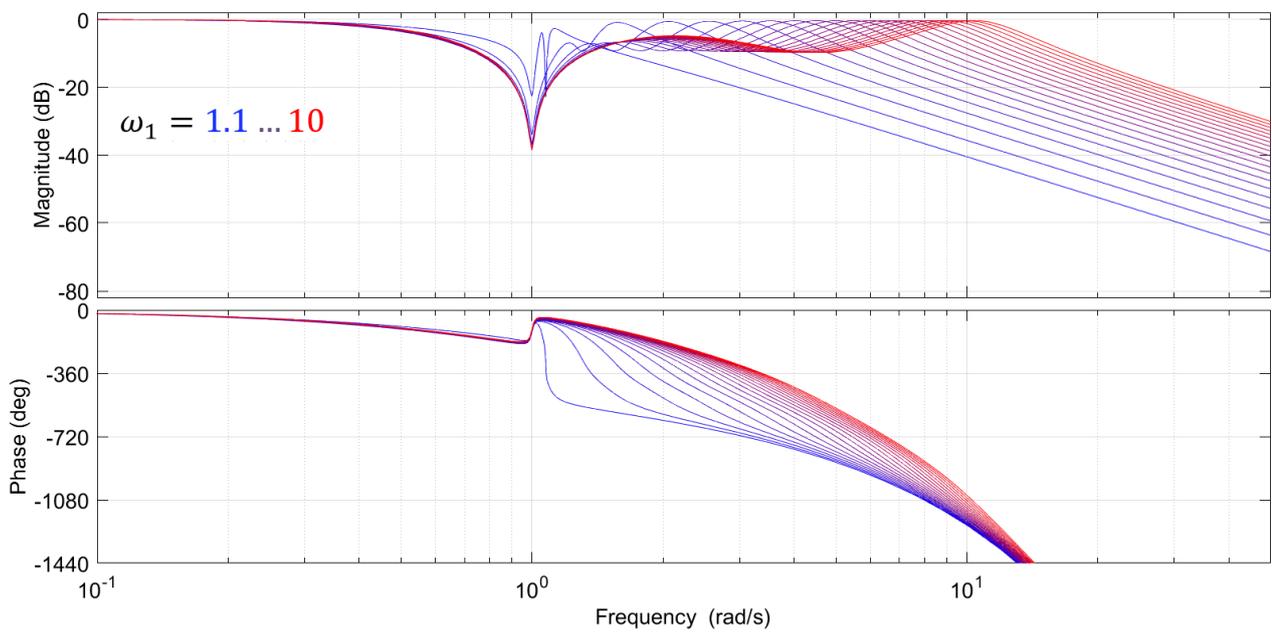


FIGURE 5.22 – Diagramme de Bode de  $T(\omega_1) = \frac{C(\omega_1)P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$  obtenu pour  $\omega_1 \in \Omega_\star$

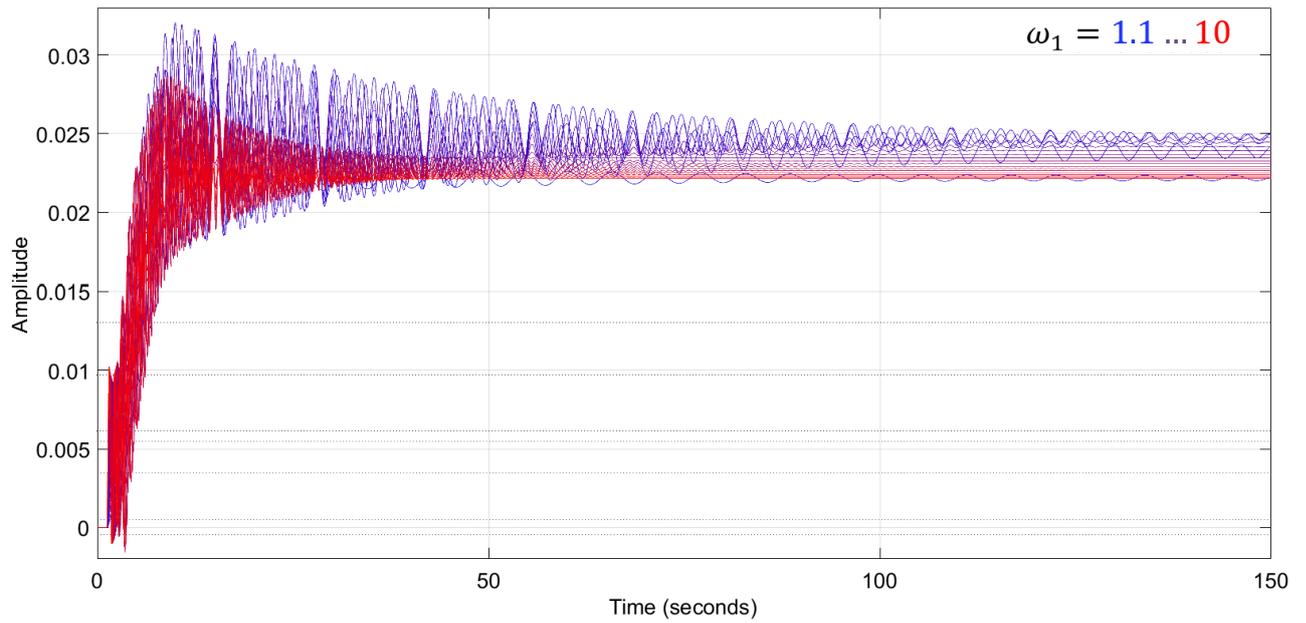


FIGURE 5.23 – Réponse à un échelon de  $\frac{P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$  obtenue pour  $\omega_1 \in \Omega_*$

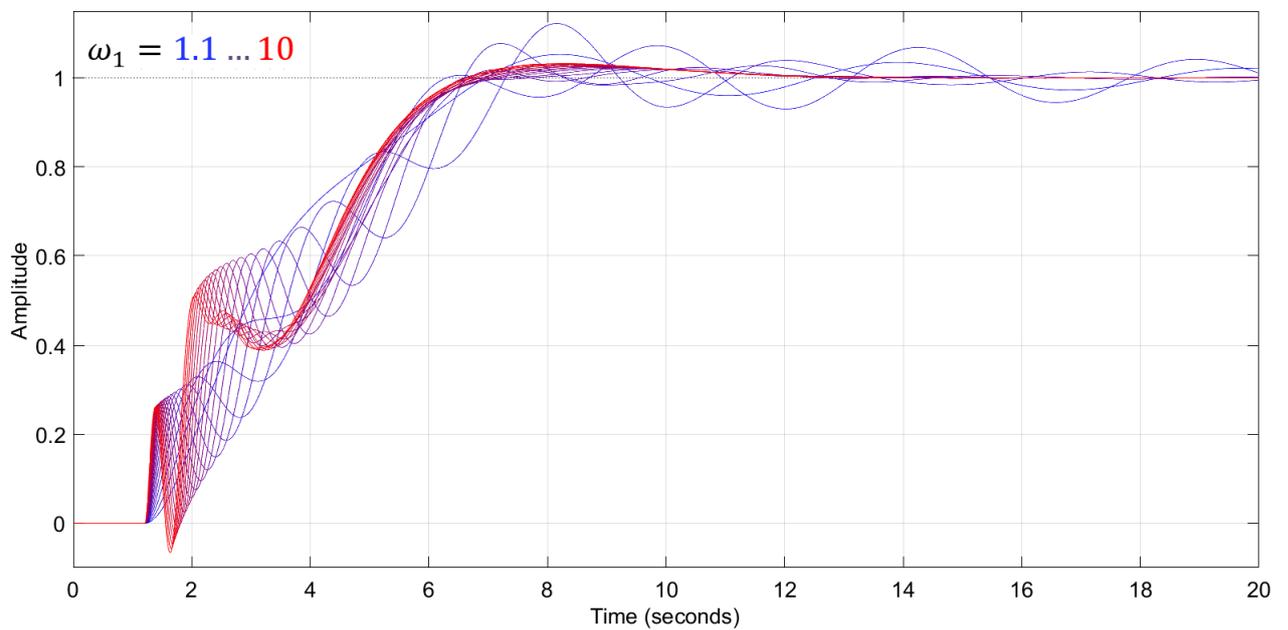


FIGURE 5.24 – Réponse à un échelon de  $\frac{C(\omega_1)P(\omega_1)}{1+C(\omega_1)P(\omega_1)}$  obtenue pour  $\omega_1 \in \Omega_*$

## 5.4 Conclusion

Dans ce chapitre, nous avons étudié la possibilité d'étendre les méthodes développées dans le cadre des systèmes SISO sans retard à une classe de systèmes SISO à retard. Nous avons montré qu'il était possible de concevoir un correcteur pour ces systèmes à retard en calculant un correcteur d'un système sans retard et nous avons exhibé des correcteurs explicites pour ces systèmes à retard. Nous avons ensuite proposé un algorithme de conception de correcteurs  $H_\infty$  par Loop-Shaping pour des systèmes à retard (et qui conserve certaines marges de stabilité). Celui-ci se base sur le réglage d'une pondération scalaire afin de garder la norme du projecteur  $\Pi$  du système sans retard constante pour des variations de paramètres, cette même norme constituant un indicateur servant à caractériser la robustesse de la boucle fermée du système avec retard.

Bien que le résultat de cette méthode ne soit pas encore bien compris théoriquement, nous remarquons que celui-ci permet d'obtenir de très bons correcteurs adaptatifs pour des systèmes à retard en pratique dans le cadre de nos applications. Il propose une méthode simple pour régler les paramètres du système bouclé.

D'une manière générale, la décomposition additive permet d'étendre les correcteurs symboliques obtenus dans le cas des systèmes linéaires SISO d'ordre inférieur à 4 à des systèmes plus complexes comme certaines classes de systèmes à retard.

# Chapitre 6

## Conclusion et perspectives

### 6.1 Résumé des contributions

Dans ce mémoire, nous avons étudié la possibilité de réaliser de manière effective des correcteurs  $H_\infty$  pour des systèmes linéaires continus invariants dans le temps, de dimension finie d'ordre faible, à une entrée et une sortie, à paramètres non fixés, c'est-à-dire de concevoir des correcteurs robustes dont les coefficients dépendent des paramètres non fixés du système (par exemple, des paramètres de modèle). Nous nous sommes concentrés sur le problème de commande  $H_\infty$  par Loop-Shaping appliqué aux systèmes SISO dont la résolution passe par le calcul des solutions d'une ERA.

Dans un premier temps, nous avons montré comment paramétrer les solutions d'une ERA en fonction d'une des inconnues. Cette dernière inconnue doit vérifier une équation polynomiale  $p = 0$  de degré  $2^n$  en toute généralité ( $2^{n-1}$  si on fixe  $b_0$ ). Lorsque le degré de ce polynôme  $p$  est inférieur ou égal à 4, nous avons explicité ses racines exprimées par radicaux à partir desquelles nous avons déduit des solutions explicites de l'ERA en fonction des paramètres du système. Dans le cas contraire, une étape numérique est alors nécessaire, c'est-à-dire une étape de particularisation des paramètres dans certaines zones régulières. Nous avons ensuite vu comment caractériser l'unique solution définie positive d'une telle équation suivant des racines du polynôme  $p$ . Nous en avons alors déduit des formes explicites des solutions définies positives pour des ERA de taille inférieure ou égale à 4. Ce premier travail permet déjà d'établir une commande  $H_2$  pour la classe considérée de systèmes à paramètres inconnus.

Nous avons ensuite exploité ce travail sur les équations de Riccati algébriques afin de concevoir des correcteurs  $H_\infty$  de systèmes d'ordre faible à paramètres inconnus. Là encore, nous avons obtenu des correcteurs dépendant explicitement des paramètres d'un système si ce dernier est d'ordre inférieur ou égal à 4. Lorsque les polynômes intervenant dans le problème de commande robuste ne sont pas solvables entièrement symboliquement, nous avons proposé une méthode symbolique/numérique certifiée permettant d'exploiter le formalisme algébrique du problème. Ensuite, nous avons exploité cette formulation algébrique afin de trouver le réglage optimal d'une pondération fréquentielle scalaire lorsque d'autres paramètres du système sont inconnus.

Enfin, nous avons montré comment étendre cette approche à une classe de systèmes à retard. Pour cela, nous avons proposé de décomposer de tels systèmes à retard en une somme d'une partie sans retard (c'est-à-dire de dimension finie) et d'une partie avec retard (conte-

nant un retard distribué). Disposant d'un correcteur stabilisant la partie sans retard, nous déduisons alors un correcteur stabilisant le système avec retard. Dans le cas de systèmes dont la partie sans retard est d'ordre faible, nous disposons alors de correcteurs sous optimaux au sens du critère  $H_\infty$  pour ces systèmes à retard, correcteurs dépendant explicitement des paramètres de ces systèmes. Enfin, nous proposons une méthode de réglage de robustesse du système avec retard en réglant le critère  $H_\infty$  du système sans retard.

De manière plus générale, cette thèse a permis d'établir certaines limites théoriques et pratiques liées au calcul de correcteurs  $H_\infty$  paramétriques à l'aide de techniques de calcul formel. En effet, bien qu'il soit possible théoriquement de trouver une paramétrisation des solutions de l'ERA de systèmes d'ordre grand, il s'avère que ce n'est pas le cas en pratique. Ceci s'explique par l'explosion exponentielle de la taille des coefficients de la paramétrisation en fonction du nombre de paramètres. Sur nos machines, nous n'avons pas pu calculer de telles représentations pour des systèmes généraux d'ordre supérieur ou égal à 6. Ceci constitue une première limite pratique. De plus, si nous disposons d'une paramétrisation des solutions de cette équation suivant un polynôme  $p$ , alors les racines de ce polynôme ne peuvent s'exprimer sous forme close que si  $p$  est de degré inférieur ou égal à 4 (du fait de la théorie de Galois). Ceci constitue une seconde limite théorique. Cette limitation se retrouve à la fois lors du calcul de solutions explicites de l'ERA intervenant dans le problème de commande, mais également dans le calcul du critère  $H_\infty$ .

Nous avons alors cherché à contourner ces limites en intégrant une étape numérique dans le processus. L'utilisation de ces algorithmes symboliques-numériques nécessite le calcul de paramétrisations pour des systèmes évalués en les paramètres. Ce genre de calcul reste raisonnable lorsque  $n$  est inférieur à 10. Ceci constitue une nouvelle limite pratique, bien que celle-ci soit nettement moins restrictive que les précédentes. Aussi, cette étape symbolique-numérique apporte une plus-value intéressante par rapport à des méthodes numériques classiques car elle *certifie numériquement le résultat* en donnant un encadrement certifié de ce dernier.

Par ailleurs, cette thèse a permis une compréhension plus globale du problème de commande  $H_\infty$  par Loop-Shaping. Les outils du calcul formel sont bien adaptés pour étudier la dépendance du critère  $H_\infty$  en fonction des paramètres du système. C'est ce que nous avons appelé *étudier les sensibilités du critère aux variations paramétriques*. Les outils du calcul formel permettent, par exemple, de tracer des graphiques représentant l'évolution du critère suivant les paramètres et ce avec une précision aussi grande que souhaitée (au prix d'un coût de calcul augmentant suivant la précision).

Enfin, nous avons montré comment utiliser cette formulation algébrique afin de régler une pondération Loop-Shaping scalaire. Pour cela, nous avons étudié une certaine courbe. Cette dernière est constituée d'un nombre fini de branches, dont une seule nous intéresse. Nous avons proposé une méthode permettant d'identifier cette branche. Ensuite, nous avons proposé deux méthodes permettant de régler le paramètre de réglage introduit par le Loop-Shaping afin de rester sur la branche adéquate. La première consiste à trouver une expression explicite exacte de cette branche. La seconde consiste à approcher localement la branche par un développement de Taylor ou de Puiseux. Le choix de la méthode est imposé par l'ordre et le nombre de paramètres du système. Cette méthodologie de conception systématique de pondération optimale est un atout supplémentaire de nos méthodes.

## 6.2 Perspectives

Cette thèse a permis de proposer de nombreuses idées et de soulever de nombreuses questions qui n'ont pu être étudiées par faute de temps.

La question de l'extension de nos méthodes au cas MIMO n'a pas été étudiée dans sa globalité. En effet, nous nous sommes concentrés sur les systèmes SISO sans transfert direct afin de travailler avec des représentations de systèmes simples. Il est établi qu'il existe des équivalences de nos travaux dans le cas MIMO comme le montre [KYAH09]. Il reste cependant à établir une paramétrisation des solutions de l'ERA dans le cas MIMO de manière systématique. Notons que, dans [BRB<sup>+</sup>18], nous calculons la solution définie positive d'une ERA à paramètres inconnus associée à un système MIMO particulier provenant de l'estimation de la position d'un drone. Ce résultat montre la faisabilité du calcul explicite de la solution définie positive d'une ERA pour certains systèmes MIMO avec des paramètres de modèle.

L'utilisation de racines par radicaux pour représenter les solutions est aussi un point à étudier. En effet, ce genre de représentation n'est pas certifié d'un point de vue numérique. En revanche, celles-ci sont faciles d'utilisation et rapidement embarquables dans un calculateur. Une comparaison des temps de calcul d'un algorithme certifié d'isolation de racines (par exemple, la méthode des signes de Descartes) et de l'évaluation de nos racines par radicaux sur calculateur est à réaliser. En environnement de simulation MATLAB, nous avons remarqué que l'évaluation de nos solutions explicites par radicaux du problème  $H_\infty$  fournissait un résultat plus rapidement que les algorithmes itératifs de MATLAB. Cette rapidité de calcul est accrue lorsque le nombre de paramètres est faible. Il reste cependant à faire des comparaisons sérieuses entre les différentes méthodes numériques de calcul de solutions du problème  $H_\infty$ .

L'étude de la résolution de l'ERA pour des systèmes sans paramètres à l'aide de l'analyse par intervalle sera développée dans un travail ultérieur.

Nous avons aussi expliqué comment régler de manière optimale des pondérations Loop-Shaping scalaires de type gain. La question du réglage de pondérations plus complexes (comportant des dynamiques additionnelles telles que des passe-bas, intégrateurs, réjecteurs, etc.) reste dans le cas général sans réponse. En effet, plus les pondérations considérées sont d'ordre élevé, plus l'ordre du système fictif augmente et donc plus les systèmes polynomiaux sous-jacents se complexifient. Quelques tests montrèrent que l'utilisation de pondérations du type  $W = w \frac{1+t_h s}{t_h s}$  introduisaient des minima locaux dans le critère  $H_\infty$  pour des systèmes comportant une inertie et un mode mécanique. L'étude du réglage optimal de telles pondérations semble alors être bien plus difficile que pour des pondérations scalaires. Cependant, elles présentent un réel intérêt pratique dans le cadre des viseurs où nous souhaitons introduire des comportements intégrateurs en basse fréquence dans le correcteur afin de rejeter les couples perturbateurs.

De même que pour l'ajout de comportements intégrateurs en basse fréquence dans le correcteur, le problème d'ajout d'effets filtrants dans le correcteur n'a pas été abordé. En pratique, ceci est réalisé car le modèle est peu connu en haute fréquence, et qu'il est nécessaire de concevoir un asservissement robuste à ces incertitudes de modèle.

Dans le dernier chapitre, nous avons étudié l'utilisation de décompositions additives afin d'obtenir des correcteurs pour des certaines classes de systèmes à retard. L'utilisation de décompositions semble être la clé afin d'augmenter la complexité des modélisations que l'on peut considérer par nos méthodes. En effet, à l'aide de telles décompositions, nous scindons

les problèmes en plusieurs sous-problèmes qu'on est capable de traiter. Il reste à examiner d'autres types de décomposition, telles que, par exemple, la décomposition multiplicative.

La comparaison de nos techniques à la commande LPV reste aussi à explorer. Quelques tests sur un exemple pratique montrent que ces deux approches restent similaires, et présentent toutes deux avantages et inconvénients. La méthode LPV permet de traiter des systèmes d'ordre plus élevé, tandis que notre approche permet de prendre en compte des variations paramétriques très larges.

Une grande partie de cette thèse est consacrée aux équations de Riccati algébriques. Ces équations interviennent notamment dans le filtrage invariant. Dans le cadre d'une collaboration avec *Safran Tech*, nous avons alors cherché à étendre nos méthodes au filtrage invariant pour application à la navigation inertielle. Ceci a donné lieu à l'article [BRB<sup>+</sup>18]. Cet article présente la résolution symbolique d'une ERA de taille  $6 \times 6$  dont les solutions sont assez creuses structurellement. Notons aussi que ces mêmes ERA interviennent dans d'autres problèmes de l'automatique, comme par exemple le traitement du signal ou la théorie des jeux différentiels, domaines dans lesquels ce travail de thèse pourrait aussi avoir un intérêt.

Dans cette thèse, nous avons étudié un problème de commande  $H_\infty$  appliqué à des systèmes invariants dans le temps. L'adaptation de nos méthodes de commande  $H_\infty$  paramétriques à la commande adaptative reste possible pour des systèmes variant lentement dans le temps (par rapport aux dynamiques du système). Ceci est le cas pour certains systèmes développés par Safran, où les paramètres changent suivant la température ou l'usure du système. Dans le cas de systèmes variant rapidement dans le temps, il reste intéressant d'étendre ces techniques symboliques aux *équations de Riccati différentielles*. Ceci constitue une nouvelle perspective de recherche future.

Pour finir, une des perspectives les plus attrayantes de ce travail reste son applicabilité à la commande adaptative. En effet, les correcteurs obtenus étant facilement implémentables sur un simple calculateur, la question se pose d'identifier plus précisément le modèle de notre système en temps réel afin de pouvoir adapter le correcteur aux variations paramétriques du système. Sur des exemples, nous avons vu que si l'on disposait d'une identification parfaite des paramètres du modèle, alors la commande adaptative obtenue était très performante. Cependant, pour cela, il est nécessaire de concevoir une stratégie d'identification des paramètres du modèle continu. Nous pouvons citer, par exemple, [FSR03], [LZ05] ou [GWY08] qui proposent des méthodes différentes pour étudier ce problème. L'identification des systèmes est un problème important, et demande un travail en profondeur. Notamment, il n'existe pas de méthodes génériques qui permettent de traiter l'identification de systèmes généraux. Ces méthodes diffèrent si le système est en boucle ouverte ou fermée, s'il comporte un retard ou un intégrateur pur, s'il est proche de la non-observabilité, si les signaux sont échantillonnés à pas variable ou non, etc. De plus, la qualité de l'identification varie en fonction du rapport signal sur bruit et de la fréquence d'échantillonnage. Tous ces paramètres montrent qu'il est nécessaire d'acquérir un certain savoir-faire sur ces méthodes avant de pouvoir adapter notre méthodologie de commande à de la commande adaptative. Notamment, cette thématique de recherche sur l'identification en ligne de modèles de viseurs fera l'objet d'une prochaine thèse portée par Safran et Inria Lille Nord-Europe.

# Annexe A

## Systèmes linéaires continus invariants dans le temps

Nous présentons ici quelques rappels sur les systèmes linéaires continus de dimension finie définis par des équations différentielles linéaires à coefficients constants, sur la stabilité, l'observabilité et la commandabilité.

### A.1 Définition d'un système linéaire

**Définition A.1.** Un *système linéaire continu à temps invariant*  $G$  (abrégé LTI, pour Linear Time Invariant) est un opérateur de convolution entre un vecteur d'entrée  $u(t) \in \mathbb{R}^m$  à support dans  $\mathbb{R}^+$  (c'est-à-dire  $u(t) = 0$  pour tout  $t < 0$ ) et un vecteur de sortie  $y(t) \in \mathbb{R}^p$ , c'est-à-dire

$$G : u \mapsto y = H * u,$$

où  $(h * u)(t) = \int_0^t h(s) u(t-s) ds$ . Nous parlerons aussi de *transfert* entre un vecteur d'entrées  $u \in \mathbb{R}^m$  et un vecteur de sorties  $y \in \mathbb{R}^p$ .

Soit  $s := \frac{d}{dt}$  l'opérateur de dérivation, ou variable de Laplace, la *transformée de Laplace* étant définie par  $\widehat{f}(s) = \int_0^\infty e^{-st} f(t) dt$ , où  $s \in \mathbb{C}$  est proprement choisi. Il existe deux grandes manières de représenter un système linéaire à temps continu.

**Définition A.2.** Soit un système LTI  $G$  de vecteur d'entrées  $u \in \mathbb{R}^m$  et de vecteur de sorties  $y \in \mathbb{R}^p$ . Une *représentation d'état* de  $G$  est une représentation de la forme

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \\ x(t_0) = x_0, \end{cases} \quad (\text{A.1})$$

où  $x \in \mathbb{R}^n$  est appelé *vecteur d'état*,  $t_0 \in \mathbb{R}$  est l'*instant initial* (pris généralement nul),  $x_0$  l'*état initial* du système (pris généralement nul).

**Définition A.3.** Soit un système LTI  $G$ , de vecteur d'entrées  $u(t) \in \mathbb{R}^m$  et de vecteur de sorties  $y(t) \in \mathbb{R}^p$ . Nous désignons par  $U(s)$  (resp.  $Y(s)$ ) la transformée de Laplace du vecteur  $u$  (resp.  $y$ ).  $G$  peut se modéliser par une matrice de transfert  $G(s) \in \mathbb{R}(s)^{m \times p}$  telle que :

$$Y(s) = G(s)U(s).$$

**Remarque A.1.** (Système invariant dans le temps) La définition d'un système LTI se généralise aux *systèmes linéaires variants dans le temps* (abrégé LTV, pour Linear Time Varying). Nous ne rentrerons pas ici dans les détails. Notons cependant que, dans ce cas, les matrices  $A$ ,  $B$ ,  $C$  et  $D$  d'une représentation d'état d'un système  $G$  dépendent alors explicitement du temps.

## A.2 Stabilité d'un système linéaire

**Définition A.4.** Soit un système LTI  $G$  de vecteur d'entrées  $u \in \mathbb{R}^m$  et de vecteur de sorties  $y \in \mathbb{R}^p$ . Nous dirons que  $G$  est *stable d'entrée bornée et de sortie bornée*, abrégé BIBO stable, si pour vecteur d'entrée borné, c'est-à-dire,  $u \in L^\infty(\mathbb{R}_+)^m$ , alors le vecteur de sortie est borné, c'est-à-dire  $y \in L^\infty(\mathbb{R}_+)^p$ . Nous dirons alors, par abus de langage, que  $G$  est *stable*.

**Remarque A.2.** Il existe d'autres notions de stabilité : exponentielle, asymptotique, etc. Nous choisissons de ne pas les détailler ici. Notons simplement que pour un système linéaire commandable et observable (voir Section A.3), celles-ci sont toutes équivalentes. Pour plus de détails, voir, par exemple, [Kha96, CZ12].

Afin de donner une caractérisation d'un système linéaire stable, nous rappelons la notion de *matrice stable*.

**Définition A.5.** (matrice stable) Une matrice  $A \in \mathbb{R}^{n \times n}$  est dite *stable* si ses valeurs propres sont à partie réelle strictement négatives.

Nous avons alors la propriété suivante.

**Définition A.6.** (système stable) Un système linéaire invariant dans le temps, de représentation d'état  $(A, B, C, D)$  (resp. de fonction de transfert  $G$ ) est *stable* si  $A$  est stable (resp. si  $G$  a tous ses pôles à partie réelle strictement négative).

## A.3 Commandabilité et observabilité

La commandabilité se définit comme suit.

**Définition A.7** (Chap. 8.1). [CD12] Un système (A.1) est dit *commandable* si pour tout vecteur d'état initial  $x_i$  et tout vecteur d'état final  $x_f$ , il existe une commande  $u$  permettant au vecteur d'état d'atteindre l'état final  $x_f$  à partir de l'état initial  $x_i$  en un temps fini.

Cette définition étant difficilement exploitable mathématiquement, un critère de commandabilité est le suivant.

**Proposition A.1.** [Kal63] Un système linéaire de représentation d'état  $(A, B, C, D)$  est commandable si et seulement si :

$$\text{rang} \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix} = n.$$

D'une manière générale, l'observabilité se définit comme suit.

**Définition A.8** (Chap. 8.1). [CD12] Un système (A.1) est dit *observable* si l'observation de ses entrées et sorties pendant un intervalle de temps fini  $[t_i, t_f]$  permet de déterminer l'état initial du système, et donc, par intégration de l'équation d'état, de connaître le vecteur  $x(t)$  à tout instant  $t$  de l'intervalle  $[t_i, t_f]$ .

Cette définition étant aussi difficilement exploitable mathématiquement, un critère d'observabilité est le suivant.

**Proposition A.2.** [Kal63] Un système linéaire de représentation d'état  $(A, B, C, D)$  est observable si et seulement si :

$$\text{rang} \begin{pmatrix} C \\ C A \\ \dots \\ C A^{n-1} \end{pmatrix} = n.$$

## A.4 Stabilisabilité et détectabilité

Les notions de stabilisabilité et de détectabilité sont des conditions moins restrictives que les notions de commandabilité et d'observabilité. Dans un premier temps, définissons celles-ci.

**Définition A.9.** • Supposons que (A.1) ne soit pas commandable. Alors, après un changement de base du vecteur d'état, celui-ci peut être scindé en faisant apparaître un état commandable et un état non commandable :

$$\begin{cases} \begin{pmatrix} \dot{x}_c(t) \\ \dot{x}_{nc}(t) \end{pmatrix} = \begin{pmatrix} A_c & A_{12} \\ 0 & A_{nc} \end{pmatrix} \begin{pmatrix} x_c(t) \\ x_{nc}(t) \end{pmatrix} + \begin{pmatrix} B_c \\ 0 \end{pmatrix} u(t), \\ y(t) = (C_c \ C_{nc}) \begin{pmatrix} x_c(t) \\ x_{nc}(t) \end{pmatrix} + D u(t), \\ x(t_0) = x_0. \end{cases} \quad (\text{A.2})$$

Le système (A.2) est dit *stabilisable* si sa partie non commandable est stable, c'est-à-dire si  $A_{nc}$  est stable.

• Supposons que (A.1) ne soit pas observable. Alors, après un changement de base du vecteur d'état, celui-ci peut être scindé en faisant apparaître un état observable et un état non observable :

$$\begin{cases} \begin{pmatrix} \dot{x}_o(t) \\ \dot{x}_{no}(t) \end{pmatrix} = \begin{pmatrix} A_o & 0 \\ A_{21} & A_{no} \end{pmatrix} \begin{pmatrix} x_o(t) \\ x_{no}(t) \end{pmatrix} + \begin{pmatrix} B_o \\ B_{no} \end{pmatrix} u(t), \\ y(t) = (C_o \ 0) \begin{pmatrix} x_o(t) \\ x_{no}(t) \end{pmatrix} + D u(t), \\ x(t_0) = x_0. \end{cases} \quad (\text{A.3})$$

Le système (A.3) est dit *détectable* si sa partie non observable est stable, c'est-à-dire si  $A_{no}$  est stable.

**Corollaire A.1.** Un système commandable est stabilisable, et un système observable est détectable.

Nous avons alors le théorème suivant.

**Théorème A.1.** [CD12, Chap. 8.7] Le système (A.1) est stabilisable si et seulement s'il existe une matrice  $F \in \mathbb{R}^{m \times n}$  telle que  $A + B F$  est stable. Le système (A.1) est détectable si et seulement s'il existe une matrice  $K \in \mathbb{R}^{n \times p}$  telle que  $A + K C$  est stable.

## A.5 Formes compagnes et matrices de passage

Soit un système LTI de dimension  $n \in \mathbb{N}$ , SISO, strictement propre, commandable et observable, noté  $G$ . Nous désignons par  $(A, B, C)$  une *représentation sous forme compagne de commande* de  $G$  de transfert direct nul, c'est-à-dire une représentation d'état de la forme

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad B := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^{n \times 1},$$

$$C := (c_0 \quad c_1 \quad \dots \quad c_{n-1}) \in \mathbb{R}^{1 \times n},$$

et par  $(A_o, B_o, C_o)$  une *représentation sous forme compagne d'observation* de  $G$  de transfert direct nul, c'est-à-dire une représentation d'état de la forme :

$$A_o := \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad B_o := \begin{pmatrix} c_0 \\ \vdots \\ c_{n-2} \\ c_{n-1} \end{pmatrix} \in \mathbb{R}^{n \times 1},$$

$$C_o := (0 \quad 0 \quad \dots \quad 0 \quad 1) \in \mathbb{R}^{1 \times n}.$$

Par ailleurs, notons que  $A_o = A^T$ ,  $B_o = C^T$ ,  $C_o = B^T$ . Nous avons alors le théorème suivant.

**Théorème A.2.** La matrice suivante

$$\mathcal{P} = (\mathcal{P}_1 \quad \dots \quad \mathcal{P}_i \quad \dots \quad \mathcal{P}_n),$$

avec

$$\left\{ \begin{array}{l} \mathcal{P}_n := C^T, \\ \mathcal{P}_{n-1} := (A^T + a_{n-1} I_n) C^T, \\ \vdots \\ \mathcal{P}_2 := \left( (A^T)^{n-2} + a_{n-1} (A^T)^{n-3} + \dots + a_2 I_n \right) C^T, \\ \mathcal{P}_1 := \left( (A^T)^{n-1} + a_{n-1} (A^T)^{n-2} + \dots + a_1 I_n \right) C^T, \end{array} \right.$$

est inversible.

Notons  $\mathcal{Q}$  son inverse.  $\mathcal{Q}$  est une matrice de passage permettant de passer de la forme compagne d'observation à la forme compagne de commande de  $G$ . De plus, celle-ci possède les propriétés suivantes

$$\left\{ \begin{array}{l} A = \mathcal{Q} A_o \mathcal{Q}^{-1} = \mathcal{Q} A^T \mathcal{Q}^{-1}, \\ B = \mathcal{Q} B_o = \mathcal{Q} C^T, \\ C = C_o \mathcal{Q}^{-1} = B^T \mathcal{Q}^{-1}, \end{array} \right.$$

ou, de manière équivalente :

$$\begin{cases} A = \mathcal{P}^{-1} A^T \mathcal{P}, \\ B = \mathcal{P}^{-1} C^T, \\ C = B^T \mathcal{P}. \end{cases} \quad (\text{A.4})$$

*Démonstration.* Soit  $\mathcal{P} := (\mathcal{P}_1 \ \cdots \ \mathcal{P}_i \ \cdots \ \mathcal{P}_n) \in \mathbb{R}^{n \times n}$ . Remarquons que :

$$\underbrace{\begin{pmatrix} C (A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I_n) \\ C (A^{n-2} + a_{n-1} A^{n-3} + \cdots + a_2 I_n) \\ \vdots \\ C (A + a_{n-1} I_n) \\ C \end{pmatrix}}_{\mathcal{P}^T} = \underbrace{\begin{pmatrix} 1 & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \\ 0 & 1 & a_{n-1} & \cdots & a_3 & a_2 \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1} & \vdots \\ 0 & 0 & \cdots & 0 & 1 & a_{n-1} \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}}_{T_n} \times \underbrace{\begin{pmatrix} C A^{n-1} \\ C A^{n-2} \\ \vdots \\ C A \\ C \end{pmatrix}}_{\mathcal{O}_i}.$$

Le système est supposé observable, donc  $\mathcal{O}_i$  est de rang plein (corollaire immédiat de la Proposition A.2). De plus,  $T_n$  est de rang plein, donc  $\mathcal{P}$  est inversible.

Soit  $\widetilde{\mathcal{P}} \in \mathbb{R}^{n \times n}$  une matrice symétrique inversible vérifiant (A.4). Notons :

$$\widetilde{\mathcal{P}} = \left( \widetilde{\mathcal{P}}_1 \ \cdots \ \widetilde{\mathcal{P}}_i \ \cdots \ \widetilde{\mathcal{P}}_n \right).$$

Avec l'égalité

$$C = B^T \widetilde{\mathcal{P}}^T = (0 \ \cdots \ 0 \ 1) \left( \widetilde{\mathcal{P}}_1 \ \cdots \ \widetilde{\mathcal{P}}_{n-1} \ \widetilde{\mathcal{P}}_n \right)^T = \widetilde{\mathcal{P}}_n^T,$$

nous obtenons  $\widetilde{\mathcal{P}}_n := C^T$ . De plus,  $\widetilde{\mathcal{P}}$  doit aussi vérifier  $\widetilde{\mathcal{P}} A = A^T \widetilde{\mathcal{P}}$ , ce que nous pouvons réécrire :

$$\left( \widetilde{\mathcal{P}}_1 \ \cdots \ \widetilde{\mathcal{P}}_n \right) A = \left( A^T \widetilde{\mathcal{P}}_1 \ \cdots \ A^T \widetilde{\mathcal{P}}_n \right).$$

Décomposons la matrice  $A$  de la manière suivante :

$$A = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & & \cdots & & 0 \end{pmatrix}}_{I_s} + \underbrace{\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ -a_0 & \cdots & -a_{n-1} \end{pmatrix}}_{A_b}.$$

Nous avons alors

$$\widetilde{\mathcal{P}} I_s = \left( 0 \ \widetilde{\mathcal{P}}_1 \ \cdots \ \widetilde{\mathcal{P}}_i \ \cdots \ \widetilde{\mathcal{P}}_{n-1} \right),$$

et

$$\widetilde{\mathcal{P}} A_b = \left( -a_0 \widetilde{\mathcal{P}}_n \ \cdots \ -a_i \widetilde{\mathcal{P}}_n \ \cdots \ -a_{n-1} \widetilde{\mathcal{P}}_n \right),$$

ce qui amène à :

$$\left( 0 \ \widetilde{\mathcal{P}}_1 \ \cdots \ \widetilde{\mathcal{P}}_{n-1} \right) + \left( -a_0 \widetilde{\mathcal{P}}_n \ \cdots \ -a_{n-1} \widetilde{\mathcal{P}}_n \right) = \left( A^T \widetilde{\mathcal{P}}_1 \ \cdots \ A^T \widetilde{\mathcal{P}}_n \right). \quad (\text{A.5})$$

Par conséquent, nous avons

$$\begin{cases} \widetilde{\mathcal{P}}_n = C^T, \\ \widetilde{\mathcal{P}}_{n-1} = (A^T + a_{n-1} I_n) \mathcal{P}_n = (A^T + a_{n-1} I_n) C^T, \\ \widetilde{\mathcal{P}}_{n-2} = A^T \mathcal{P}_{n-1} + a_{n-2} \mathcal{P}_n = ((A^T)^2 + a_{n-1} A^T + a_{n-2} I_n) C^T, \\ \vdots \\ \widetilde{\mathcal{P}}_1 = ((A^T)^{n-1} + a_{n-1} (A^T)^{n-2} + \cdots + a_1 I_n) C^T. \end{cases}$$

et donc  $\widetilde{\mathcal{P}} = \mathcal{P}$ , ce qui termine la preuve. □

**Remarque A.3.** (A.5) montre que  $A$  est solution de l'équation caractéristique suivante :

$$A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I_n = 0.$$

# Annexe B

## Normes $H_2$ et $H_\infty$

Cette annexe a pour but de rappeler des notions sur les normes de systèmes linéaires continus et invariants dans le temps utilisés dans ce mémoire. Afin de clarifier le propos, nous introduisons les notions permettant de définir ces normes.

### B.1 Norme euclidienne

Soient  $n \in \mathbb{N}$  et un vecteur  $v \in \mathbb{R}^n$ . La *norme euclidienne* du vecteur  $v$  est alors définie par

$$\|v\| := \sqrt{v^T v},$$

où  $v^T$  dénote le transposé de  $v$ .

### B.2 Normes matricielles

Soit  $A \in \mathbb{R}^{n \times n}$ . De la norme euclidienne précédente, on déduit la norme matricielle suivante :

$$\|A\|_2 := \sup_{0 \neq v \in \mathbb{R}^n} \frac{\|A v\|}{\|v\|}.$$

### B.3 Normes de signaux

Notons  $\mathbb{R}_+ := ]0, +\infty[$ . On appelle *signal* toute fonction de  $L_2(\mathbb{R}_+)$  définie comme l'espace des fonctions mesurables au sens de Lebesgue et de carré intégrables, c'est-à-dire :

$$L_2(\mathbb{R}_+) := \left\{ f : \mathbb{R}_+ \rightarrow \mathbb{R}^n \text{ mesurables au sens de Lebesgue } \left| \int_0^\infty \|f(t)\|^2 dt < \infty \right. \right\}.$$

$L_2(\mathbb{R}_+)$  est un espace de Hilbert pour le *produit scalaire* défini par :

$$\forall x, y \in L_2(\mathbb{R}_+), \quad \langle x, y \rangle := \int_0^{+\infty} x(t) y(t) dt.$$

En particulier, la norme  $f \in L_2(\mathbb{R}_+)$  est définie par :

$$\|f\|_2 := \left( \int_0^{+\infty} \|f(t)\|^2 dt \right)^{1/2}.$$

Soit  $\mathcal{L}(f)$  la transformée de Laplace de  $f$  (voir, par exemple, [BK13, Chap. 2.4.1]). L'identité de Parseval (voir, par exemple, [BK13, Chap. 12.3.1]) donne :

$$\|f\|_2 := \frac{1}{2\pi} \left( \int_{-\infty}^{+\infty} \|\mathcal{L}(f)(i\omega)\|^2 d\omega \right)^{1/2}.$$

Rappelons que l'on appelle *valeur singulière* d'une matrice  $M \in \mathbb{R}^{m \times p}$  toute racine carrée d'une valeur propre de  $M^* M$  (où  $M^* = M^T$  dans le cas d'une matrice  $M$  réelle). Notons alors par  $\sigma_{\max}(M)$  la *valeur singulière maximale* d'une matrice  $M$ . Si  $s \mapsto \mathcal{L}(f)(s)$  est analytique bornée dans  $\mathbb{C}_+ = \{s \in \mathbb{C} \mid \Re(s) > 0\}$ , alors on peut définir la norme suivante :

$$\|\mathcal{L}(f)\|_\infty := \sup_{s \in \mathbb{C}_+} \sigma_{\max}(\mathcal{L}(f)(s)).$$

De plus, nous avons le résultat suivant [ZDG96, Chap. 4] :

$$\|\mathcal{L}(f)\|_\infty = \operatorname{ess\,sup}_{\omega \in \mathbb{R}} \sigma_{\max}(\mathcal{L}(f)(i\omega)).$$

## B.4 Normes de systèmes

Soit  $G$  un système linéaire, continu, invariant dans le temps, de représentation d'état supposée sans transfert direct ( $A, B, C$ ) (c'est-à-dire  $D = 0$ ), et de fonction de transfert :

$$G(s) = C (sI - A)^{-1} B.$$

### B.4.1 Espace $H_2(\mathbb{C}_+)$ et norme $H_2$

**Définition B.1.** Nous définissons par  $H_2(\mathbb{C}_+)$  l'espace de Hilbert formé des fonctions holomorphes dans  $\mathbb{C}_+$ , bornées pour la norme :

$$\|G\|_2 := \left( \frac{1}{2\pi} \int_{-\infty}^{+\infty} \operatorname{Tr}(G^*(i\omega) G(i\omega)) d\omega \right)^{1/2}.$$

Par abus de langage, nous parlerons de *norme  $H_2$* .

**Remarque B.1.** Dans le cas SISO, la norme  $H_2$  de  $G$  représente l'énergie en sortie du système lorsque l'entrée du système est une impulsion de type Dirac ou un bruit blanc vérifiant  $U(i\omega)U^*(i\omega) = I$  (c'est-à-dire de densité spectrale uniforme).

**Remarque B.2. Calcul numérique de la norme  $H_2$ .** Supposons que  $A$  est stable et  $D$  est nulle (dans un soucis de clarté). Notons que

$$\mathcal{L}(G(i s)) = C e^{At} B,$$

ce qui permet de déduire :

$$\int_0^{+\infty} B^T e^{tA^T} C^T C e^{tA} B dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G^*(i\omega) G(i\omega) d\omega.$$

Notons que  $Q := \int_0^{+\infty} e^{tA^T} C^T C e^{tA} dt$  est l'unique solution de l'équation de Lyapunov suivante :

$$A^T Q + Q A + C^T C = 0. \tag{B.1}$$

En passant à la trace, nous obtenons alors

$$\|G\|_2 := (\text{Tr}(B^T Q B))^{1/2},$$

où  $Q$  est solution de (B.1). Ainsi, résoudre (B.1), qui peut se faire à l'aide de des techniques d'algèbre linéaire [TVTB14], permet de déduire simplement la norme  $H_2$  d'un système linéaire continu, invariant dans le temps.

### B.4.2 Espace $H_\infty(\mathbb{C}_+)$ et norme $H_\infty$

**Définition B.2.** [Zam81, ZDG96] Nous définissons par  $H_\infty(\mathbb{C}_+)$ , aussi noté  $H_\infty$  dans ce mémoire, l'algèbre de Hardy formée par les fonctions holomorphes dans  $\mathbb{C}_+ := \{s \in \mathbb{C} \mid \text{Re}(s) > 0\}$  bornées pour la norme

$$\|G\|_\infty = \sup_{s \in \mathbb{C}_+} (\sigma_{\max}(G(s))) = \text{ess sup}_{\omega \in \mathbb{R}} (\sigma_{\max}(G(i\omega))) = \sup_{U \in H_2(\mathbb{C}_+) \neq 0} \frac{\|Y\|_2}{\|U\|_2},$$

où  $\sigma_{\max}(M)$  dénote la plus grande valeur singulière d'une matrice  $M$ . Par abus de langage, nous parlerons de *norme*  $H_\infty$ .

**Remarque B.3.** Dans le cas des systèmes SISO, la norme  $H_\infty$  de  $G$  représente le gain maximal de sa réponse fréquentielle  $G(i\omega)$ . Cette norme est particulièrement utile pour caractériser des comportements critiques (c'est-à-dire le pire des comportements possibles).

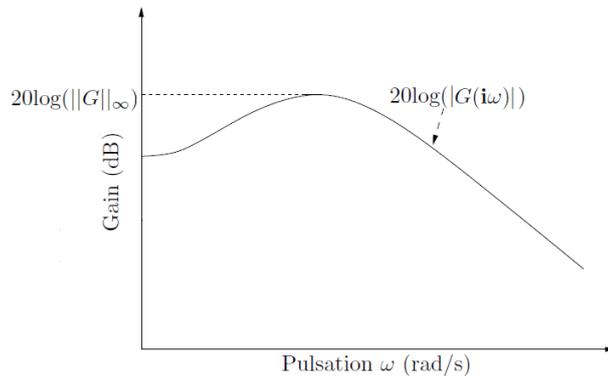


FIGURE B.1 – Norme  $H_\infty$  d'une fonction de transfert  $G$  dans le cas SISO

### Calcul numérique

Soit  $\gamma \in \mathbb{R}$ . Construisons la matrice Hamiltonienne suivante :

$$H(\gamma) := \begin{pmatrix} A & \gamma^{-1} B B^T \\ -\gamma^{-1} C^T C & -A^T \end{pmatrix}.$$

[DF00, ACA<sup>+</sup>99] montre que  $\|G\|_\infty < \gamma$  si et seulement si  $H(\gamma)$  n'a pas de valeur propre sur l'axe imaginaire. L'algorithme suivant permet alors d'approximer la norme  $H_\infty$  de  $G$  :

1. Choisir un intervalle grossier  $[\gamma_{\min}, \gamma_{\max}]$ .
2. Calculer le spectre de  $H(\gamma)$  pour  $\gamma = \frac{1}{2}(\gamma_{\min} + \gamma_{\max})$ .
3. Si  $H(\gamma)$  n'a pas de valeurs propres sur l'axe imaginaire, alors  $\gamma$  est trop grand. On peut alors raffiner l'intervalle en revenant à l'étape 1 et avec le nouveau encadrement  $[\gamma_{\min}, \gamma]$ . Sinon,  $\gamma$  est trop petit et on peut raffiner l'intervalle en revenant à l'étape 1 et avec l'encadrement  $[\gamma, \gamma_{\max}]$ .

### Calcul symbolique

Dans le cas SISO, pour une fonction de transfert  $G$ , calculer sa norme  $H_\infty$  revient à calculer les zéros du numérateur de  $\frac{d|G(i\omega)|^2}{d\omega}$ , tout en s'assurant que le dénominateur ne s'annule pas sur ces zéros. Le travail est donc limité au calcul de variations d'une fonction. Nous avons suivi cette méthode pour calculer la norme  $H_\infty$  d'une fonction de transfert donnée dans la Proposition 5.2.

Dans le cas MIMO, pour une matrice de transfert  $G$ , [KS06] propose une méthode basée sur le calcul du discriminant en  $s^2$  de :

$$\Phi_\gamma(s) = \det(\gamma^2 I - G^*(s)G(s)).$$

Pour la notion de discriminant, nous renvoyons à la Section 3.1.10 du Chapitre 3. La norme  $H_\infty$  de  $G$  est alors donnée par un des cas suivants :

- une racine de ce discriminant  $\text{disc}(\Phi_\gamma(s), s^2)$ ,
- la valeur singulière maximale de  $G$  en  $s = 0$ ,
- la valeur singulière maximale de  $G$  en  $s = i\infty$ .

Pour plus de détails, se référer à [KS06].

**Remarque B.4.** Les deux méthodes symboliques énoncées ici sont équivalentes dans le cas des systèmes SISO. Ceci peut se comprendre en se rappelant que le calcul d'un discriminant revient à calculer le lieu d'annulation des zéros communs d'un polynôme et de sa dérivée.

### Inégalités classiques

Pour finir, nous rappelons quelques propriétés classiques de la norme  $H_\infty$  que l'on utilise dans ce mémoire de thèse :

1.  $\|H_1 H_2\|_\infty \leq \|H_1\|_\infty \|H_2\|_\infty, \quad H_1, H_2 \in RH_\infty.$
2.  $\left\| \begin{pmatrix} H_1 & H_2 \\ H_3 & H_4 \end{pmatrix} \right\|_\infty \leq \gamma \quad \Rightarrow \quad \begin{cases} \|H_1\|_\infty \leq \gamma, \\ \|H_2\|_\infty \leq \gamma, \\ \|H_3\|_\infty \leq \gamma, \\ \|H_4\|_\infty \leq \gamma. \end{cases} \quad H_1, H_2, H_3, H_4 \in RH_\infty$

Pour plus de détails, voir, par exemple, [ACA<sup>+</sup>99] ou [Fey13].

# Annexe C

## Matrices définies positives

**Définition C.1.** Soit  $M \in \mathbb{R}^{n \times n}$  telle que  $M^T = M$ . Alors,  $M$  est dite *définie positive* si :

$$\forall x \in \mathbb{R}^n, \quad x \neq 0, \quad x^T M x > 0.$$

Nous noterons alors  $M > 0$ .

Nous rappelons maintenant quelques caractéristiques des matrices définies positives.

**Proposition C.1.** (*Caractérisation des matrices définies positives*) [Bha09] Soit  $M \in \mathbb{R}^{n \times n}$  telle que  $M^T = M$ . Les propositions suivantes sont équivalentes :

- $M > 0$ .
- Toutes les valeurs propres de  $M$  sont réelles strictement positives.
- La forme bilinéaire symétrique suivante

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^T M y \end{aligned}$$

est un produit scalaire sur  $\mathbb{R}^n$ .

- Il existe une matrice  $N \in \mathbb{R}^{n \times n}$  *invertible* telle que :

$$M = N^T N.$$

### C.1 Produits de matrices définies positives

**Proposition C.2.** Soit  $A, B \in \mathbb{R}^{n \times n}$  des matrices définies positives. Alors les valeurs propres du produit  $AB$  sont réelles strictement positives.

*Démonstration.* La Proposition C.1 montre que si  $A$  (resp.,  $B$ ) est une matrice réelle définie positive, alors il existe une matrice non-singulière  $M \in \mathbb{R}^{n \times n}$  (resp.  $N$ ) telle que  $A = M^T M$  (resp.,  $B = N^T N$ ). Par conséquent, nous avons :

$$AB = M^T M N^T N, \quad (AB)^T = (N^T N)(M^T M).$$

Si nous notons  $L := M^T M N^T$ , alors nous avons  $AB = LN$ . Comme  $N$  est non-singulière, nous avons  $L = (AB)N^{-1}$ , ce qui permet d'obtenir :

$$N(AB)N^{-1} = NL = NM^T M N^T = (MN^T)^T (MN^T).$$

En notant  $P := M N^T$ , nous avons alors :

$$N (A B) N^{-1} = P^T P.$$

De plus, en notant que  $\det(M) \neq 0$  et  $\det(N) \neq 0$ , nous avons :

$$\det(P) = \det(M N^T) = \det(M) \det(N^T) = \det(M) \det(N) \neq 0.$$

Par conséquent,  $P$  n'est pas singulière, donc inversible. Ainsi,  $A B$  est similaire à la matrice définie positive  $(M N^T)^T (M N^T)$ . En particulier, les valeurs propres de  $A B$  sont les mêmes que celles de  $(M N^T)^T (M N^T)$ , qui sont réelles strictement positives, ce qui conclut la preuve.  $\square$

En particulier, le résultat précédent peut s'appliquer aux matrices  $X$  et  $Y$  solutions de  $\mathcal{R} = 0$  et  $\tilde{\mathcal{R}} = 0$ .

**Remarque C.1.** Notons que :

$$(A B)^T = N^T N M^T M \neq A B.$$

Ainsi,  $A B$  n'est pas nécessairement symétrique, et donc n'est pas nécessairement définie positive, bien que  $A$  et  $B$  le soient. Exhibons un contre exemple. Soient les matrices définies positives suivantes :

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} > 0,$$

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} > 0.$$

Alors, nous avons :

$$A B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} \neq (A B)^T.$$

# Annexe D

## Racines par radicaux de polynômes de degré inférieur ou égal à 4

Dans cette annexe, nous explicitons les racines par radicaux de polynômes de degré inférieur ou égal à 4. Aussi, nous identifions celle correspondant à la racine réelle maximale selon les valeurs des paramètres.

Ces racines par radicaux font intervenir des puissances non entières de nombres complexes. Afin de garder une unicité des notations, nous adopterons la notation suivante

$$x \in \mathbb{C}, n \in \mathbb{Q}, \quad x^{1/n} = \exp\left(\frac{1}{n} (\ln(|x|) + i \arg(x))\right),$$

où  $\arg(x)$  est l'argument de  $x$  tel que  $-\pi \leq \arg(x) \leq \pi$ . Dans la littérature,  $x^{1/n}$  définie ci-dessus est appelée *racine n<sup>ème</sup> principale* de  $x$ . Voir [BC09, p.26] pour plus de détails.

### D.1 Polynômes de degré 2

Soit  $\nu_0, \nu_1 \in \mathbb{R}$ . Nous considérons le polynôme de degré 2 suivant :

$$P_2 := X^2 + \nu_1 X + \nu_0. \tag{D.1}$$

#### D.1.1 Racines par radicaux

Soit  $\epsilon \in \{-1, 1\}$ . Les racines de  $P_2$  s'écrivent :

$$x(\epsilon) := \frac{-\nu_1 + \epsilon \sqrt{\nu_1^2 - 4\nu_0}}{2}. \tag{D.2}$$

#### D.1.2 Racine réelle maximale

Notons  $\Delta_2 := \nu_1^2 - 4\nu_0$  le discriminant de  $P_2$  en  $X$ . Suivant le signe de  $\Delta_2$ , le nombre de racines réelles de  $P_2$  change.

Cas  $\Delta_2 > 0$  :  $P_2$  possède deux racines réelles. Trivialement, la plus grande est :

$$x_{\max} := \frac{-\nu_1 + \sqrt{\nu_1^2 - 4\nu_0}}{2}.$$

Cas  $\Delta_2 = 0$  :  $P_2$  possède une racine réelle double :

$$x_{\max} := \frac{-\nu_1}{2}.$$

Cas  $\Delta_2 < 0$  :  $P_2$  possède deux racines complexes et aucune racine réelle.

L'analyse de ces différents cas peut se compléter par une représentation graphique de la variété discriminante  $\Delta_2 = 0$ . Dans la Figure D.1, le cas  $\Delta_2 < 0$  est représenté en rouge, le cas  $\Delta_2 \geq 0$  en bleu, et le cas critique  $\Delta_2 = 0$  en noir.

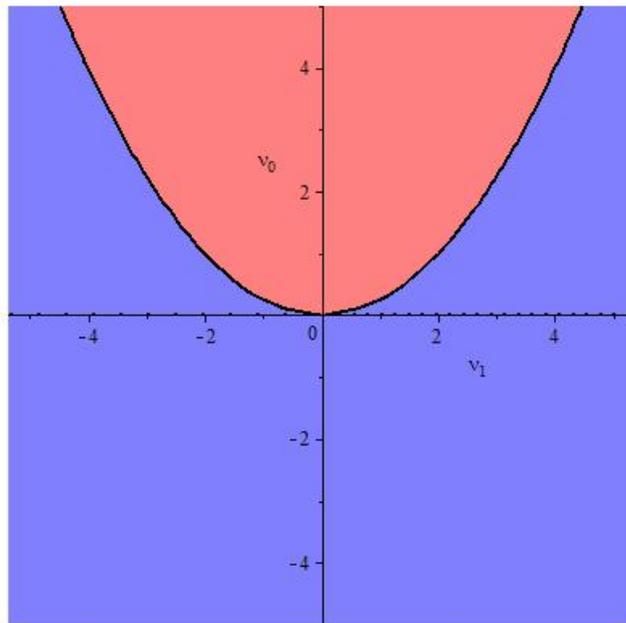


FIGURE D.1 – Variété discriminante  $\delta_2 = 0$  de  $P_2$

Nous retiendrons simplement le résultat suivant.

**Proposition D.1.** Si  $P_2$  admet au moins une racine réelle, alors sa racine réelle maximale est donnée par :

$$x_{\max} := \frac{-\nu_1 + \sqrt{\nu_1^2 - 4\nu_0}}{2}.$$

## D.2 Polynômes de degré 3

Soient  $\nu_i \in \mathbb{R}$  pour  $i = 0, 1, 2$ . Nous considérons le polynôme de degré 3 suivant :

$$P_3 := X^3 + \nu_2 X^2 + \nu_1 X + \nu_0. \quad (\text{D.3})$$

### D.2.1 Racines par radicaux

En appliquant le changement de variable  $X = Y - \nu_2/3$  à (D.3), nous éliminons le terme en  $X^2$  de  $P_3$  et nous trouvons le polynôme suivant

$$Q_3 := Y^3 + \mu_1 Y + \mu_0, \quad (\text{D.4})$$

où :

$$\begin{cases} \mu_1 := -\frac{1}{3} \nu_2^2 + \nu_1, \\ \mu_0 := \frac{2}{27} \nu_2^3 - \frac{1}{3} \nu_1 \nu_2 + \nu_0. \end{cases}$$

Nous proposons de reprendre la méthode de Lagrange afin d'exprimer les racines par radicaux de  $P_3$ . Celle-ci permet de réduire la résolution de  $Q_3 = 0$  à la recherche de racines d'un polynôme de degré 2. Cependant, cette analogie n'est possible sous une condition particulière, ce qui nous amène à distinguer deux cas.

• **Cas  $\mu_1 = 0$ .**

Supposons  $\mu_1 = 0$  et notons  $j := -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ . Nous cherchons alors les racines de :

$$Q_3 := Y^3 + \mu_0.$$

Leur construction est triviale : les racines s'expriment via la racine cubique de  $\mu_0$  et les racines cubiques de  $(-1)$ , c'est-à-dire  $(-1)$ ,  $-j$  et  $-j^2$ . Ceci est résumé dans la proposition suivante.

**Proposition D.2.** Supposons  $\nu_1 = \frac{1}{3} \nu_2^2$  et notons  $\mu_0 := \frac{2}{27} \nu_2^3 - \frac{1}{3} \nu_1 \nu_2 + \nu_0$ . Les racines de  $P_3 := X^3 + \nu_2 X^2 + \nu_1 X + \nu_0$ , notées  $x_1, x_2, x_3$ , s'expriment de la façon suivante :

$$\begin{cases} x_1 = (-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ x_2 = j^2 (-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ x_3 = j (-\mu_0)^{1/3} - \frac{\nu_2}{3}. \end{cases} \quad (\text{D.5})$$

• **Cas  $\mu_1 \neq 0$ .**

Supposons maintenant  $\mu_1 \neq 0$ . Notons  $y_1, y_2, y_3$  les racines de  $Q_3$ . A partir des coefficients de  $Q_3$ , nous pouvons écrire les *sommes de Newton* suivantes :

$$\begin{cases} y_1 + y_2 + y_3 = 0, \\ y_1 y_2 + y_2 y_3 + y_1 y_3 = \mu_1, \\ y_1 y_2 y_3 = -\mu_0. \end{cases}$$

Notons que  $j$  vérifie  $1 + j + j^2 = 0$ . Notons alors :

$$\begin{cases} s_1 := y_1 + y_2 j + y_3 j^2, \\ s_2 := y_1 + y_3 j + y_2 j^2. \end{cases}$$

A l'aide de ces propriétés, après quelques calculs ([Mut80], p. 58), nous trouvons

$$\begin{cases} s_1 s_2 = -3 \mu_1, \\ s_1^3 + s_2^3 = -27 \mu_0, \end{cases} \quad (\text{D.6})$$

soit :

$$\begin{cases} s_1^3 s_2^3 = -27 \mu_1^3, \\ s_1^3 + s_2^3 = -27 \mu_0. \end{cases} \quad (\text{D.7})$$

En d'autres termes, (D.7) définit le produit et la somme de deux quantités  $s_1$  et  $s_2$ . On peut donc dire que ces deux quantités sont solutions d'une équation du second degré vérifiant ces deux propriétés (la somme et le produit des racines se retrouvant directement dans les coefficients).  $s_1^3$  et  $s_2^3$  sont alors solutions de l'équation suivante :

$$E_z := Z^2 + 27 \mu_0 Z - 27 \mu_1^3 = 0. \quad (\text{D.8})$$

Notant  $\Delta_3 := 27(4\mu_1^3 + 27\mu_0^2)$ , les racines de  $E_z$  s'écrivent alors :

$$\begin{cases} s_1^3 = \frac{-27 \mu_0 + \sqrt{\Delta_3}}{2}, \\ s_2^3 = \frac{-27 \mu_0 - \sqrt{\Delta_3}}{2}. \end{cases}$$

Soulignons aussi que si  $s_1$  est une racine cubique de  $s_1^3$ , c'est-à-dire une racine de  $E_z$ , alors (voir (D.6)) :

$$s_2 = -\frac{3\mu_1}{s_1}.$$

Notons alors  $\alpha_3$  la racine cubique de  $s_1^3$  suivante :

$$\alpha_3 := \left( \frac{-27 \mu_0 + \sqrt{\Delta_3}}{2} \right)^{1/3}.$$

Il ne nous reste alors plus qu'à résoudre le système linéaire suivant :

$$\begin{cases} y_1 + y_2 + y_3 = 0, \\ y_1 + y_2 j + y_3 j^2 = \alpha_3, \\ y_1 + y_3 j + y_2 j^2 = -\frac{3\mu_1}{\alpha_3}. \end{cases}$$

On obtient alors les trois racines de  $Q_3 = Y^3 + \mu_1 Y + \mu_0$  :

$$\begin{cases} y_1 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right), \\ y_2 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right) j^2, \\ y_3 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right) j. \end{cases}$$

Des expressions de  $y_1, y_2, y_3$ , on en déduit les expressions des racines de  $P_3$ , notées  $x_1, x_2, x_3$  :

$$\begin{cases} x_1 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right) - \frac{\nu_2}{3}, \\ x_2 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right) j^2 - \frac{\nu_2}{3}, \\ x_3 = \frac{1}{3} \left( \alpha_3 - \frac{3\mu_1}{\alpha_3} \right) j - \frac{\nu_2}{3}. \end{cases} \quad (\text{D.9})$$

### Récapitulatif

En regroupant les résultats des deux cas précédents, nous obtenons la proposition suivante.

**Proposition D.3.** Notons par  $(x_1, x_2, x_3)$  les expressions par radicaux des racines de  $P_3 := X^3 + \nu_2 X^2 + \nu_1 X + \nu_0$ . Notons aussi :

$$\begin{cases} \mu_1 := -\frac{1}{3}\nu_2^2 + \nu_1, \\ \mu_0 := \frac{2}{27}\nu_2^3 - \frac{1}{3}\nu_1\nu_2 + \nu_0, \\ \Delta_3 = 27(4\mu_1^3 + 27\mu_0^2), \\ \alpha_3 := \left(\frac{-27\mu_0 + \sqrt{\Delta_3}}{2}\right)^{\frac{1}{3}}. \end{cases}$$

Alors, nous avons :

$$\begin{aligned} \text{Si } \mu_1 = 0, & \quad \begin{cases} x_1 = (-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ x_2 = j^2(-\mu_0)^{1/3} - \frac{\nu_2}{3}, \\ x_3 = j(-\mu_0)^{1/3} - \frac{\nu_2}{3}. \end{cases} \\ \text{Si } \mu_1 \neq 0, & \quad \begin{cases} x_1 = \frac{1}{3}\left(\alpha_3 - \frac{3\mu_1}{\alpha_3}\right) - \frac{\nu_2}{3}, \\ x_2 = \frac{1}{3}\left(\alpha_3 - \frac{3\mu_1}{\alpha_3}\right)j^2 - \frac{\nu_2}{3}, \\ x_3 = \frac{1}{3}\left(\alpha_3 - \frac{3\mu_1}{\alpha_3}\right)j - \frac{\nu_2}{3}. \end{cases} \end{aligned}$$

**Remarque D.1.** Supposons  $\mu_1 = 0$ . Alors,  $\Delta_3 = (27\mu_0)^2$  et :

$$\alpha_3 = \left(\frac{-27\mu_0 + |27\mu_0|}{2}\right)^{1/3}.$$

Si  $\mu_0 = \frac{2}{27}\nu_2^3 - \frac{1}{3}\nu_1\nu_2 + \nu_0 < 0$ , alors  $\alpha_3 = 3(-\mu_0)^{\frac{1}{3}} \neq 0$  et (D.9) est bien défini. Nous remarquons alors que les valeurs des racines données par (D.5) ou (D.9) coïncident. Dans ce cas, nous avons alors deux représentations possibles pour les racines de  $P_3$ , (D.5) étant plus “légère” dans ce cas.

Par contre, si  $\mu_0 > 0$ , alors  $\alpha_3 = 0$  et (D.9) n’est plus défini. Dans ce cas, seul (D.5) est défini.

La variété  $\mu_1 = 0$  est très importante en pratique. En effet, à l’approche de cette variété, les expressions par radicaux (D.9) tendent vers une forme indéterminée de la forme “zéro divisé par zéro”. Cela pourra poser des problèmes d’évaluation des racines lorsque la précision de calcul disponible est limitée.

Par exemple, pour  $\epsilon \in \mathbb{R}$ , posons :

$$\begin{cases} \nu_0 = 1, \\ \nu_2 = 1, \\ \nu_1 = \frac{1}{3}\nu_2^2 + \epsilon. \end{cases}$$

Pour  $\epsilon = 0$ , les valeurs de paramètres choisies sont exactement sur la variété  $\mu_1 = 0$ . En choisissant  $\epsilon = 1/100$ , l'évaluation des expressions (D.9) avec 12 chiffres après la virgule donne des résultats proches des véritables racines (obtenues par une isolation certifiée) à  $10^{-4}$  près, tandis que l'évaluation de (D.5) n'est précise qu'à  $10^{-1}$  près. En revanche, pour  $\epsilon = 1/1000$ , l'évaluation numérique de (D.9) n'est précise qu'à  $10^{-1}$  près tandis que (D.5) est précise à  $10^{-3}$  près. Pour  $\epsilon = 1/100000$ , nous constatons que l'évaluation numérique de (D.9) diverge. Il est donc difficile d'avoir une bonne précision de calcul en évaluant des expressions par radicaux à l'approche de cette variété. Il pourra donc être préférable d'utiliser une technique d'isolation de racines, voir par exemple [CA76].

$\epsilon$	Eval. num. des formes closes (D.9) (cas $\mu_1 \neq 0$ )	Eval. num. des formes closes (D.5) (cas $\mu_1 = 0$ )	Isolation numérique certifiée
1/100	-1.31630193883	-1.32083202280	-1.31631186149
1/1000	-1.33300000000	-1.32083202280	-1.32038047536
1/100000	$-\infty$	-1.32083202280	-1.32038047536

TABLE D.1 – Test

**Remarque D.2.** [Zwi02] donne une expression des racines de  $P_3$  faisant intervenir des fonctions trigonométriques. Pour  $\mu_1 \neq 0$ , celles-ci s'écrivent :

$$\tilde{x}_i := 2 \sqrt{\frac{-\mu_1}{3}} \cos \left( \frac{1}{3} \arccos \left( \frac{3\mu_0}{2\mu_1} \sqrt{\frac{-3}{\mu_1}} \right) - \frac{2(i-1)\pi}{3} \right) - \frac{\nu_2}{3}, \quad i = 1, 2, 3. \quad (\text{D.10})$$

**Remarque D.3.** Lorsque  $P_3$  possède une seule racine réelle et si  $\mu_1 \neq 0$ , [Hol02] montre que l'on peut exprimer cette racine réelle à l'aide de fonctions hyperboliques :

$$x_r := \begin{cases} -2 \operatorname{sgn}(\mu_0) \sqrt{\frac{-\mu_1}{3}} \cosh \left( \frac{1}{3} \operatorname{arccosh} \left( -\frac{3|\mu_0|}{2\nu_1} \sqrt{\frac{-3}{\mu_1}} \right) \right) - \frac{\nu_2}{3}, & \text{si } \begin{cases} \mu_1 < 0, \\ 4\mu_1^3 + 27\mu_0^2 > 0, \end{cases} \\ -2 \sqrt{\frac{\mu_1}{3}} \sinh \left( \frac{1}{3} \operatorname{arcsinh} \left( \frac{3\mu_0}{2\mu_1} \sqrt{\frac{3}{\mu_1}} \right) \right) - \frac{\nu_2}{3}, & \text{si } \mu_1 > 0. \end{cases} \quad (\text{D.11})$$

## D.2.2 Racine réelle maximale

### Avec des radicaux

Soit  $P_3 := X^3 + \nu_2 X^2 + \nu_1 X + \nu_0$ . Notons par  $x_1, x_2, x_3$  les racines de  $P_3$  conformément à la Proposition D.3. Notons  $\Delta$  son discriminant divisé par  $(-27)$  :

$$\Delta := \Delta_3/27 = \nu_0^2 + 2 \frac{2\nu_2^2 - 9\nu_1}{27} \nu_2 \nu_0 + \frac{4\nu_1 - \nu_2^2}{27} \nu_1^2.$$

Les racines de  $\Delta$  en  $\nu_0$ , qui nous serviront par la suite, s'écrivent :

$$\delta_1 := \frac{\nu_2}{3} \left( -\frac{2\nu_2^2}{9} + \nu_1 \right) + \frac{2}{27} \sqrt{(\nu_2^2 - 3\nu_1)^3}, \quad \delta_2 := \frac{\nu_2}{3} \left( -\frac{2\nu_2^2}{9} + \nu_1 \right) - \frac{2}{27} \sqrt{(\nu_2^2 - 3\nu_1)^3}. \quad (\text{D.12})$$

Nous avons alors la proposition suivante.

**Proposition D.4.** La racine  $x_1$  est la racine réelle maximale de  $P_3$  si et seulement si :

$$\begin{cases} \mu_1 = 0, \\ \mu_0 \leq 0, \end{cases} \quad \text{ou} \quad \begin{cases} \nu_2 \in \mathbb{R}, \\ \nu_1 \in \left] \frac{\nu_2^2}{3}; \infty \right[ , \\ \nu_0 \in \mathbb{R}, \end{cases} \quad \text{ou} \quad \begin{cases} \nu_2 \in \mathbb{R}, \\ \nu_1 \in \left] -\infty; \frac{\nu_2^2}{3} \right[ , \\ \nu_0 \in \left] -\infty; \delta_1 \right]. \end{cases}$$

La racine  $x_3$  est la racine réelle maximale de  $P_3$  si et seulement si :

$$\begin{cases} \mu_1 = 0, \\ \mu_0 \geq 0, \end{cases} \quad \text{ou} \quad \begin{cases} \nu_2 \in \mathbb{R}, \\ \nu_1 \in \left] -\infty; \frac{\nu_2^2}{3} \right[ , \\ \nu_0 \in [\delta_1; \infty[. \end{cases}$$

*Démonstration.* Nous cherchons quelle expression par radicaux, parmi celles de la Proposition D.3, correspond à la racine réelle maximale du polynôme  $P_3$ . Cette expression est susceptible de varier suivant les valeurs des paramètres  $\nu_0$ ,  $\nu_1$  et  $\nu_2$ . Si tel est le cas, nous aimerions savoir pour quelle condition cette expression change.

Nous proposons alors d'identifier le lieu des racines sur lequel deux racines de  $P_3$  sont égales. Ce lieu s'appelle la variété discriminante de  $P_3$  (voir Section 3.1.10). Celle-ci, notée  $\Delta = 0$ , sépare l'espace des paramètres en cellules (c'est-à-dire en semi-algébriques de l'espace des paramètres) dans lesquelles le nombre de racines réelles de  $P_3$  est constant (voir Section 3.1.10). Disposant des frontières de ces cellules, une isolation numérique certifiée des racines en un point donné de chaque cellule permet de dénombrer le nombre de racines réelles dans ladite cellule. Noter que nous pourrions aussi évaluer les expressions par radicaux des racines afin de déterminer le nombre de racines réelles, mais nous pourrions perdre l'aspect certifié. Pour plus de détails sur les variétés discriminantes, se référer à [LR07] ou [Mor08]. Pratiquement, ce genre de calcul peut se faire à l'aide de `Maple` et de la commande `CellDecomposition` du package `RootFinding[Parametric]`.

Cependant, notons que nous ne voulons pas uniquement déterminer combien de racines réelles comporte chaque cellule, mais aussi à quelle expression par radicaux correspond la racine réelle maximale. Nous devons alors prendre en compte le domaine de définition des expressions par radicaux introduites dans la Proposition D.3.

Étudier la variété discriminante de  $P_3$  permet de savoir si deux ou plus racines de  $P_3$  coïncident lorsque les paramètres  $\nu_0$ ,  $\nu_1$  et  $\nu_2$  changent, le lieu sur lesquelles elles se rencontrent étant exactement la variété discriminante. Cette propriété est uniquement liée au polynôme  $P_3$  et est indépendante de la représentation par radicaux des racines de  $P_3$  que nous avons mis en avant.

• Soit  $\mu_1 = 0$ . L'annulation du discriminant de  $P_3$  coïncide alors exactement avec l'annulation de  $\mu_0$ . On distingue alors deux cas :  $\mu_0 \leq 0$  et  $\mu_0 \geq 0$ . L'évaluation des racines (D.5) dans un point de chaque cas permet de conclure quant à l'expression de la racine maximale.

• Soit  $\mu_1 \neq 0$ . Nous pouvons alors considérer le discriminant générique de  $P_3$ , noté  $\Delta$ . Une décomposition de l'espace des paramètres suivant cette variété discriminante, et l'évaluation des racines par radicaux (D.9) en chaque point de chaque cellule permet alors de conclure quant à l'expression de la racine maximale de  $P_3$ .

Noter que lorsque  $\nu_0 = \delta_1$  ou  $\nu_0 = \delta_2$ ,  $P_3$  admet une racine multiple car on est alors sur la variété discriminante. Ainsi, pour ces valeurs de paramètres, il existe plusieurs expressions par radicaux donnant la racine multiple maximale. Ceci est indiqué dans la Proposition D.4 par des crochets fermants pour les intervalles en  $\nu_0$ .

□

**Remarque D.4.** Noter que la borne  $\nu_1 = \nu_2^3/3$  est racine du discriminant de  $\Delta_3$  en  $\nu_2$ . Cette borne correspond exactement au cas particulier où les expressions (D.9) ne sont pas définies.

De la Proposition D.4, on en déduit le corollaire suivant.

**Corollaire D.1.** Si les racines de  $P_3$  sont toutes réelles, alors l'expression par radicaux de la solution maximale est  $x_1$ .

### Avec fonctions trigonométriques

Supposons  $\mu_1 \neq 0$ . Conformément à la Remarque D.2, notons par  $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$  les expressions par fonctions trigonométriques des racines de  $P_3$ . Nous avons alors le résultat suivant.

**Proposition D.5.** La forme close  $\tilde{x}_1$  représente la réelle maximale de  $P_3$  si et seulement si :

$$\nu_2 \in \mathbb{R}, \quad \nu_1 < \frac{\nu_2^2}{3}, \quad \nu_0 \leq \delta_1.$$

La forme close  $\tilde{x}_2$  est la seule racine réelle de  $P_3$  (d'où la réelle maximale) si et seulement si :

$$\nu_2 \in \mathbb{R}, \quad \nu_1 > \frac{\nu_2^2}{3}, \quad \nu_0 \in \mathbb{R}.$$

La forme close  $\tilde{x}_3$  est la seule racine réelle de  $P_3$  (d'où la réelle maximale) si et seulement si :

$$\nu_2 \in \mathbb{R}, \quad \nu_1 < \frac{\nu_2^2}{3}, \quad \nu_0 \geq \delta_1.$$

Lorsque  $3\nu_1 - \nu_2^2 = 0$ , il y a une seule racine réelle (voir Proposition D.2).

*Démonstration.* Comme pour la Proposition D.4, la preuve se base sur le calcul d'une variété discriminante de  $P_3$ . □

**Remarque D.5.** Lorsque les racines de  $P_3$  sont toutes réelles, l'expression par radicaux de la solution maximale est  $\tilde{x}_1$ . Les racines de  $P_3$  sont toutes réelles si et seulement si :

$$\nu_2 \in \mathbb{R}, \quad \nu_1 < \frac{\nu_2^2}{3}, \quad \delta_2 \leq \nu_0 \leq \delta_1.$$

## D.3 Polynômes de degré 4

Pour  $i \in \{0, 1, 2, 3\}$ , soient  $\nu_i \in \mathbb{R}$ . Considérons le polynôme de degré 4 suivant :

$$P_4 := X^4 + \nu_3 X^3 + \nu_2 X^2 + \nu_1 X + \nu_0. \quad (\text{D.13})$$

### D.3.1 Racines par radicaux

Cette section reprend la *méthode de Ferrari* pour exprimer les racines de  $P_4$  [Mut80].

En appliquant le changement de variable  $X = Y - \frac{\nu_3}{4}$  à  $P_4$ , nous éliminons le terme en  $X^3$  et nous obtenons le polynôme suivant

$$Q_4 := Y^4 + \mu_2 Y^2 + \mu_1 Y + \mu_0, \quad (\text{D.14})$$

où :

$$\begin{cases} \mu_2 := -\frac{3}{8} \nu_3^2 + \nu_2, \\ \mu_1 := \frac{1}{8} \nu_3^3 - \frac{1}{2} \nu_2 \nu_3 + \nu_1, \\ \mu_0 := -\frac{3}{256} \nu_3^4 + \frac{1}{16} \nu_2 \nu_3^2 - \frac{1}{4} \nu_1 \nu_3 + \nu_0. \end{cases}$$

Nous allons maintenant travailler sur  $Q_4 = 0$  pour faire apparaître des termes au carré. Premièrement, de  $Q_4 = 0$ , nous avons :

$$\left(Y^2 + \frac{\mu_2}{2}\right)^2 = -\mu_1 Y - \mu_0 + \frac{1}{4} \mu_2^2. \quad (\text{D.15})$$

On aimerait faire apparaître un carré parfait dans le membre de droite. On distingue alors deux cas. Le premier correspond au cas où ce carré parfait apparaît naturellement, c'est-à-dire  $\mu_1 = 0$ . Le polynôme  $Q_4$  est alors quadratique en  $Y^2$ . Le second correspond à un cas plus général où  $\mu_1 \neq 0$ .

- **Cas  $\mu_1 = 0$ .**

Supposons  $\mu_1 = 0$ . L'équation (D.14) est un carré parfait en  $Y$ . Soient  $(\varepsilon_1, \varepsilon_2) \in \{-1, 1\}^2$ . On peut alors se ramener aux formules d'une équation quadratique afin d'obtenir le carré des racines de  $Q_4$  :

$$y^2 := \frac{1}{2} \left( \varepsilon_2 \sqrt{\mu_2^2 - 4\mu_0 - \mu_2} \right).$$

Notons  $\varepsilon := (\varepsilon_1, \varepsilon_2)$ . Les racines de  $P_4$  s'écrivent alors :

$$x(\varepsilon) := \varepsilon_1 \frac{\sqrt{2}}{2} \sqrt{\varepsilon_2 \sqrt{\mu_2^2 - 4\mu_0 - \mu_2}}.$$

- **Cas  $\mu_1 \neq 0$ .**

Supposons maintenant  $\mu_1 \neq 0$ . On peut introduire un paramètre  $u_4 \in \mathbb{C}$  afin de modifier la partie de droite de (D.15)

$$\left(Y^2 + \frac{\mu_2}{2} + u_4\right)^2 = -\mu_1 Y - \mu_0 + \frac{1}{4} \mu_2^2 + 2u_4 Y^2 + \mu_2 u_4 + u_4^2, \quad (\text{D.16})$$

l'objectif étant d'arriver à une expression de la forme, pour  $u_4 \neq 0$  :

$$\left(Y^2 + \frac{\mu_2}{2} + u_4\right)^2 = \left(\sqrt{2u_4}Y - \frac{1}{2\sqrt{2u_4}}\mu_1\right)^2. \quad (\text{D.17})$$

En développant le terme de droite de (D.17), et en égalisant avec le terme de droite de (D.16), on trouve que  $u_4$  doit satisfaire l'équation cubique en  $U$  suivante :

$$E_u := U^3 + \mu_2 U^2 + \left(\frac{1}{4}\mu_2^2 - \mu_0\right)U - \frac{1}{8}\mu_1^2 = 0. \quad (\text{D.18})$$

On peut trouver les solutions de cette équation par radicaux en se servant de la proposition D.3. Notons :

$$\begin{cases} p_4 := -\left(\frac{1}{12}\mu_2^2 + \mu_0\right), \\ q_4 := \frac{\mu_2}{3}\left(\mu_0 - \frac{\mu_2^2}{36}\right) - \frac{\mu_1^2}{8}, \\ \delta_4 := 27(4p_4^3 + 27q_4^2), \\ \alpha_4 := \left(\frac{-27q_4 + \sqrt{\delta_4}}{2}\right)^{\frac{1}{3}}. \end{cases}$$

Une solution de  $E_u$  est alors donnée par :

$$u_4 := \begin{cases} \frac{1}{3}\left(\alpha_4 - \frac{3p_4}{\alpha_4} - \mu_2\right), & \text{si } p_4 \neq 0, \\ (-q_4)^{1/3} - \frac{\mu_2}{3}, & \text{si } p_4 = 0. \end{cases}$$

Connaissant une solution  $u$  de  $E_u$ , on en déduit alors  $y$  solution de  $Q_4$  car l'on se retrouve avec des formes quadratiques. En effet, en prenant  $\varepsilon_1 \in \{-1, 1\}$ ,  $y$  est alors solution de

$$Y^2 + \frac{\mu_2}{2} + u_4 = \varepsilon_1 \left(\sqrt{2u_4}Y - \mu_1 \frac{1}{2\sqrt{2u_4}}\right),$$

que l'on peut réécrire :

$$E_y(\varepsilon_1) := Y^2 - \varepsilon_1 \sqrt{2u_4}Y + \frac{\mu_2}{2} + u_4 + \varepsilon_1 \mu_1 \frac{1}{2\sqrt{2u_4}} = 0. \quad (\text{D.19})$$

Le discriminant de  $E_y(\varepsilon_1)$  en  $Y$  est donné par

$$\Delta_y(\varepsilon_1) := -2 \left(u_4 + \mu_2 + \frac{\varepsilon_1 \mu_1}{\sqrt{2u_4}}\right), \quad (\text{D.20})$$

et, en notant  $\varepsilon_2 \in \{-1, 1\}$ , les solutions de (D.19) s'écrivent alors :

$$Y = \varepsilon_1 \frac{\sqrt{2}}{2} \sqrt{u_4} + \varepsilon_2 \frac{1}{2} \sqrt{\Delta_y(\varepsilon_1)}. \quad (\text{D.21})$$

Pour retrouver les solutions en  $X$ , il suffit alors d'appliquer le changement de variable  $X = Y - \frac{\nu_3}{4}$ . On a alors :

$$X = \varepsilon_1 \frac{\sqrt{2}}{2} \sqrt{u} + \varepsilon_2 \frac{1}{2} \sqrt{\Delta_y(\varepsilon_1)} - \frac{\nu_3}{4}. \quad (\text{D.22})$$

**Remarque D.6.** Dans le cas où  $\mu_1 = 0$ , l'équation (D.18) admet nécessairement une racine nulle (produit des racines égal à 0), ce qui pose des problèmes dans (D.17) (division par zéro). Ceci justifie la nécessité de distinguer plusieurs cas.

### Récapitulatif

En unifiant certaines notations et en regroupant l'ensemble des cas précédents, nous obtenons la proposition suivante.

**Proposition D.6.** Soit  $P_4 := X^4 + \nu_3 X^3 + \nu_2 X^2 + \nu_1 X + \nu_0$ . Notons :

$$\left\{ \begin{array}{l} \mu_2 := -\frac{3}{8} \nu_3^2 + \nu_2, \\ \mu_1 := \frac{1}{8} \nu_3^3 - \frac{1}{2} \nu_2 \nu_3 + \nu_1, \\ \mu_0 := -\frac{3}{256} \nu_3^4 + \frac{1}{16} \nu_2 \nu_3^2 - \frac{1}{4} \nu_1 \nu_3 + \nu_0, \\ p_4 := -\left(\frac{1}{12} \mu_2^2 + \mu_0\right), \\ q_4 := \frac{\mu_2}{3} \left(\mu_0 - \frac{\mu_2^2}{36}\right) - \frac{\mu_1^2}{8}, \end{array} \right. \quad \left\{ \begin{array}{l} \delta_4 := 27(4p_4^3 + 27q_4^2), \\ \alpha_4 := \left(\frac{1}{2}(\sqrt{\delta_4} - 27q_4)\right)^{\frac{1}{3}}, \\ u_4 := \begin{cases} \frac{1}{3} \left(\alpha_4 - \frac{3p_4}{\alpha_4} - \mu_2\right), & \text{si } p_4 \neq 0, \\ (-q_4)^{1/3} - \frac{\mu_2}{3}, & \text{si } p_4 = 0, \end{cases} \\ \varepsilon_1 := \pm 1, \quad \varepsilon_2 := \pm 1, \\ \Delta_4 := -\left(u_4 + \mu_2 + \varepsilon_1 \frac{\mu_1}{\sqrt{2u_4}}\right). \end{array} \right.$$

Les racines de  $P_4$  admettent les formes par radicaux suivantes :

$$x(\varepsilon_1, \varepsilon_2) := \begin{cases} \frac{\sqrt{2}}{2} (\varepsilon_1 \sqrt{u_4} + \varepsilon_2 \sqrt{\Delta_4}) - \frac{\nu_3}{4}, & \text{si } \mu_1 \neq 0, \\ \varepsilon_1 \frac{\sqrt{2}}{2} \sqrt{\varepsilon_2 \sqrt{\mu_2^2 - 4\mu_0} - \mu_2}, & \text{si } \mu_1 = 0. \end{cases}$$

### D.3.2 Racine réelle maximale

La section précédente donne des formules explicites pour les racines de  $P_4$ , polynôme de degré 4 en la variable  $X$ . Cette sous-section a pour objectif de montrer quelle expression par radicaux est la racine réelle maximale de  $P_4$  en fonction des valeurs des paramètres  $\nu_i$ .

Les racines de  $P_4$  s'écrivent  $x(\varepsilon_1, \varepsilon_2)$  et sont définies à la Proposition D.6. Notons  $\Delta_Q$  le discriminant du polynôme  $Q_4$  défini par (D.14) en la variable  $Y$ . Nous avons alors :

$$\left\{ \begin{array}{l} \Delta_Q := 256 (\mu_0^3 + \beta_2 \mu_0^2 + \beta_1 \mu_0 + \beta_0), \\ \beta_2 := -\frac{\mu_2^2}{2}, \\ \beta_1 := \frac{\mu_2}{16} (\mu_2^3 + 9\mu_1^2), \\ \beta_0 := -\frac{\mu_1^2}{64} \left(\mu_2^3 + \frac{27}{4} \mu_1^2\right). \end{array} \right.$$

Notons par  $r_i$ ,  $i \in \{1, 2, 3\}$ , les racines de  $\Delta_Q$  en  $\mu_0$ . Avec les notations

$$\left\{ \begin{array}{l} p_3 := -\frac{1}{3} \beta_2^2 + \beta_1, \\ q_3 := \frac{2}{27} \beta_2^3 - \frac{1}{3} \beta_1 \beta_2 + \beta_0, \\ \delta_3 = 27(4p_3^3 + 27q_3^2), \\ \alpha_3 := \left(\frac{-27q_3 + \sqrt{\delta_3}}{2}\right)^{1/3}, \end{array} \right.$$

la Proposition D.3 permet d'exprimer celles-ci comme suit :

$$\left\{ \begin{array}{l} r_1 = (-q_3)^{1/3} - \frac{\beta_2}{3}, \\ r_2 = j^2 (-q_3)^{1/3} - \frac{\beta_2}{3}, \text{ si } p_3 = 0, \\ r_3 = j (-q_3)^{1/3} - \frac{\beta_2}{3}, \end{array} \right. \quad \left\{ \begin{array}{l} r_1 = \frac{1}{3} \left( \alpha_3 - \frac{3p_3}{\alpha_3} - \beta_2 \right), \\ r_2 = \frac{1}{3} \left( \alpha_3 j^2 - \frac{3p_3}{\alpha_3 j} - \beta_2 \right), \text{ si } p_3 \neq 0. \\ r_3 = \frac{1}{3} \left( j \alpha_3 - \frac{3p_3}{\alpha_3 j^2} - \beta_2 \right). \end{array} \right.$$

Notons aussi, pour  $\mu_2 \leq 0$ ,  $\mu_1^* := \sqrt{\frac{-8\mu_2^3}{27}}$  et  $\mu_0^* := -\frac{\mu_2^2}{12}$ . De plus, définissons les cellules suivantes :

$$\left\{ \begin{array}{l} C_1 := \{ \mu_2 \leq 0, \quad \mu_1 \leq -\mu_1^*, \quad \mu_0 \leq r_1 \}, \\ C_2 := \{ \mu_2 \leq 0, \quad \mu_1 \in [-\mu_1^*; 0[, \quad \mu_0 \in ]-\infty; r_1] \setminus \{ \mu_0^* \} \}, \\ C_3 := \{ \mu_2 \leq 0, \quad \mu_1 \in ]0; \mu_1^*], \quad \mu_0 \in ]\mu_0^*; r_2] \}, \\ C_4 := \{ \mu_2 \geq 0, \quad \mu_1 < 0, \quad \mu_0 \leq r_1 \}, \\ C_5 := \{ \mu_2 \leq 0, \quad \mu_1 \in ]0; \mu_1^*], \quad \mu_0 \in ]-\infty; \mu_0^* \cup [r_2; r_1] \}, \\ C_6 := \{ \mu_2 \leq 0, \quad \mu_1 \geq \mu_1^*, \quad \mu_0 \leq r_1 \}, \\ C_7 := \{ \mu_2 \geq 0, \quad \mu_1 > 0, \quad \mu_0 \leq r_1 \}, \\ C_8 := \{ \mu_2 \leq 0, \quad \mu_1 \in [-\mu_1^*; 0[ \}, \\ C_9 := \{ \mu_2 \geq 0, \quad \mu_1 > 0 \}. \end{array} \right. \quad (\text{D.23})$$

Nous avons alors le résultat suivant.

**Proposition D.7.** • Supposons  $\mu_1 = 0$ . Alors,  $x(1, 1)$  est la racine réelle maximale de  $P_4$ .

• Supposons  $\mu_1 \neq 0$  et  $\mu_0 \neq \mu_0^*$ . Alors, nous avons les deux cas suivants :

–  $x(1, 1)$  est la racine réelle maximale de  $P_4$  si et seulement si :

$$\mu \in C_1 \cup C_2 \cup C_3 \cup C_4.$$

–  $x(1, -1)$  est la racine réelle maximale de  $P_4$  si et seulement si :

$$\mu \in C_5 \cup C_6 \cup C_7.$$

• Supposons  $\mu_1 \neq 0$  et  $\mu_0 = \mu_0^*$ . Alors, nous avons les deux cas suivants :

–  $x(1, -1)$  est racine réelle maximale de  $P_4$  si et seulement si :

$$\mu \in C_8 \cup C_9.$$

–  $x(1, 1)$  est racine maximale de  $P_4$  dans le cas contraire.

*Démonstration.* Dans un premier temps, nous donnons les grandes idées de la démonstration. Le discriminant de  $P_4$ , augmenté des variétés singulières  $\mu_0 = \mu_0^*$  et  $\mu_1 = 0$ , décompose l'espace des paramètres en cellules dans lesquelles le nombre de racines réelles de  $P_4$  est constant et dont les racines exprimées par radicaux sont continues. L'identification de l'expression par radicaux correspondant à la racine réelle maximale se fait en évaluant les expressions par radicaux en un point témoin de chaque cellule. Insistons sur le fait que  $\mu_0 = \mu_0^*$  et  $\mu_1 = 0$  sont des cas particuliers où les expressions par radicaux changent sur ces variétés. Ils doivent donc être traités séparément.

Pratiquement, cette démonstration est réalisée en utilisant le logiciel *Maple* et la commande *CellDecomposition* du package *RootFinding[Parametric]*. Pour le cas général, nous utilisons la commande

$$\text{CellDecomposition} ([P_4 = 0, \mu_1 \neq 0, \mu_2^2 + 12\mu_0 \neq 0], [X], [\mu_0, \mu_1, \mu_2]),$$

qui décompose l'espace des paramètres en 22 cellules. L'évaluation d'un point témoin en chacune de ces cellules nous permet alors d'identifier quelle expression par radicaux correspond à la solution réelle maximale dans la cellule considérée. Nous pouvons ensuite regrouper ces cellules en un plus petit nombre de cellules  $C_1, \dots, C_7$  définies en (D.23).

Pour le cas particulier  $\mu_2^2 + 12\mu_0 = 0$ , nous remplaçons  $\mu_0$  par  $\mu_0^*$  dans  $P_4$  pour obtenir le polynôme suivant :

$$P_4^*(X, \mu_2, \mu_1) := P_4(X, \mu_2, \mu_1, \mu_0 = \mu_0^*).$$

L'utilisation de la commande *Maple*

$$\text{CellDecomposition} ([P_4^* = 0, \mu_1 \neq 0], [X], [\mu_1, \mu_2])$$

permet de décomposer l'espace des paramètres en 6 cellules. Nous regroupons alors ces cellules en fonction de l'expression de la solution réelle maximale, ce qui donne les cellules  $C_8$  et  $C_9$ .

Pour le cas  $\mu_1 = 0$ , la démonstration est triviale car l'équation correspondante se ramène alors à une simple équation quadratique.  $\square$

Le corollaire suivant est une conséquence immédiate de la proposition précédente.

**Corollaire D.2.** Toutes les racines de  $P_4$  sont réelles si et seulement si :

$$\begin{cases} \mu_2 \in ]-\infty; 0], \\ \mu_1 \in [-\mu_1^*; \mu_1^*], \\ \mu_0 \in [r_3; r_2]. \end{cases}$$

De plus, dans ces intervalles, l'expression par radicaux de la solution réelle maximale est donnée par  $x(1, 1)$ .



# Annexe E

## Complexité algorithmique et binaire d'une RUR de $\mathcal{E}$

Le Chapitre 3 montre comment réduire l'étude de l'équation de Riccati algébrique  $\mathcal{R} = 0$  à l'étude d'un système d'équations polynomiales  $\mathcal{E}$ . Ce même chapitre montre comment résoudre  $\mathcal{E}$  à l'aide de méthodes symboliques telles que les bases de Gröbner ou les Représentations Univariées Rationnelles (RUR). Ces méthodes symboliques reposent sur des opérations symboliques. Par conséquent, le coût de calcul de celles-ci augmente lorsque la taille des matrices ou la taille des coefficients des matrices  $A, B, C$  augmentent.

Afin d'estimer précisément la difficulté de la résolution de ces systèmes, et ainsi pouvoir identifier les classes de systèmes pour lesquelles les algorithmes présentés au Chapitre 3 peuvent être appliqués, nous allons étudier la complexité du système  $\mathcal{E}$ . Plus précisément, nous sommes intéressés pour les bornes sur la taille de la représentation des solutions de la RUR. De telles bornes ont la particularité d'être intrinsèques dans le sens où elles ne dépendent que des objets algébriques/géométriques étudiés et non de l'algorithme utilisé pour les calculer.

Avant d'aller plus loin, rappelons quelques définitions et notations que nous utiliserons par la suite.

### E.1 Définitions et notations

**Définition E.1.** La *taille binaire* d'un nombre entier  $p$ , noté  $\tau_p$ , est le nombre de bits requis pour le représenter, c'est-à-dire  $\text{floor}(\log_2(p)) + 1$ , où  $\text{floor}(\bullet)$  désigne l'arrondi inférieur. Si  $p$  est rationnel, alors son taille binaire est le maximum des taille binaire de son numérateur et de son dénominateur. Si  $p$  est un polynôme, alors sa taille binaire est le maximum des taille binaire de ses coefficients.

Dans la suite, précisons aussi que nous utiliserons la notation  $O$  pour l'étude des bornes asymptotiques supérieures : si  $f$  et  $g$  sont deux fonctions de la variable réelle  $x$  et qu'il existe des constantes  $N$  et  $C$  telles que

$$\forall x > N, \quad |f(x)| \leq C |g(x)|,$$

alors nous dirons que  $f$  est dominée par  $g$  en  $+\infty$  et nous noterons  $f(x) = O(g(x))$ .

Rappelons aussi que le système résultant de l'ERA  $\mathcal{R} = 0$  d'ordre  $n$  est donné par l'ensemble de  $n$  polynômes  $\{\mathcal{B}_k\}_{0 \leq k \leq n-1}$  en les  $n$  variables  $b_k$ ,  $k = 0, \dots, n-1$ , où :

$$\begin{cases} \mathcal{B}_k := b_k^2 + 2 \sum_{m=1}^n (-1)^m b_{k-m} b_{k+m} - d_{2k}, & 1 \leq k \leq n-1, \\ d_{2k} := 2 \sum_{m=1}^n (-1)^m (a_{k-m} a_{k+m} + c_{k-m} c_{k+m}) + a_k^2 + c_k^2, & 1 \leq k \leq n-1. \end{cases} \quad (\text{E.1})$$

Rappelons que nous fixons :

$$\begin{cases} a_k := 0, b_k := 0, c_k := 0, & k \in \mathbb{Z} \setminus \llbracket 0, n \rrbracket, \\ a_n := 1, b_n := 1, c_n := 0. \end{cases}$$

Afin de calculer des bornes sur la taille des polynômes de la RUR, nous utilisons la formulation classique donnée dans [VdWAN50, Ren89] qui exprime ces polynômes comme facteurs d'un polynôme particulier appelé *polynôme  $u$ -résultant*.

## E.2 Résultant multivarié, $u$ -résultant et RUR

Le *polynôme  $u$ -résultant* est un cas particulier du *résultant multivarié* [CLO06, Chap. 3]. Nous rappelons sa définition ci-dessous. Pour cela, considérons  $f_1, \dots, f_n$  des polynômes homogènes en les variables  $x_1, \dots, x_n$ , chacun de degré  $d_i$ , donnés par :

$$f_i = \sum_{j_1 + \dots + j_n = d_i} u_{j_1, \dots, j_n}^{(i)} x_1^{j_1} \dots x_n^{j_n}.$$

De plus, notons par  $\mathbb{Z}[u]$  l'anneau des polynômes à coefficients entiers en les coefficients  $u_{j_1, \dots, j_n}^{(i)}$  des  $f_i$ . Nous notons par  $\mathbb{P}^n := \mathbb{P}^n(\mathbb{C}) = \mathbb{C}^{n+1} \setminus \{0\}$  l'espace projectif de  $\mathbb{C}^n$ . Le résultant multivarié de  $f_1, \dots, f_n$  se définit alors de la manière suivante.

**Définition E.2.** Le *résultant* de  $f_1, \dots, f_n$  est un polynôme irréductible de  $\mathbb{Z}[u]$  qui s'annule uniquement si  $\{f_1 = \dots = f_n = 0\}$  a une solution non triviale dans  $\mathbb{P}^n$ .

Le degré du résultant multivarié possède la propriété suivante.

**Théorème E.1.** [CLO06, Thm. 3.1] Le résultant de  $f_1, \dots, f_n$  est un polynôme homogène de degré  $\prod_{j \neq i} d_j$  en le coefficient  $u_{j_1, \dots, j_n}^{(i)}$ . De plus, le degré total de ce polynôme est  $\sum_{j=1}^n \prod_{i \neq j} d_i$ .

Dans le cas de polynômes  $f_1, \dots, f_n$  à coefficients dans  $\mathbb{Q}$ , le précédent théorème peut se réécrire, par spécialisation, de la manière suivante.

**Théorème E.2.** Soient  $f_1, \dots, f_n$  des polynômes homogènes en  $x_1, \dots, x_n$ , chacun de degré  $d_i$  et à coefficients rationnels de taille binaire  $\tau_i$ . Le résultant de  $f_1, \dots, f_n$  est un nombre rationnel de taille binaire  $O(\sum_{j=1}^n \tau_j \prod_{i \neq j} d_i)$ .

D'un point de vue calculatoire, le polynôme résultant apparaît comme un facteur spécifique du déterminant d'une matrice appelée *matrice de Macaulay*. Pour plus de détails sur la construction de cette matrice, ainsi que sur le calcul du résultant, voir [Mac02, VdWAN50] et leurs références.

Comme nous le mentionnons au début de cette annexe, nous pouvons utiliser le résultant pour calculer une RUR d'un ensemble de zéros de polynômes non-homogènes  $f_1, \dots, f_n$  en les variables  $x_1, \dots, x_n$ . Pour cela, nous pouvons suivre les étapes suivantes :

- Homogénéiser les polynômes  $f_1, \dots, f_n$  en introduisant une variable supplémentaire  $x_{n+1}$ .
- Ajouter au système un nouveau polynôme de la forme  $U := u_0 t + u_1 x_1 + \dots + u_n x_n$ , où  $t, u_0, \dots, u_n$  sont les nouvelles indéterminées.
- Calculer le résultant du nouveau système par rapport aux variables  $x_1, \dots, x_n, x_{n+1}$ .

Le résultant obtenu est alors un polynôme en les variables  $t, u_0, u_1, \dots, u_n$ . Il porte le nom de *polynôme  $u$ -résultant*. En notant ce polynôme par  $R(t, u_0, u_1, \dots, u_n)$  et en supposant que  $a_1 x_1 + \dots + a_n x_n$  est une forme linéaire séparante des zéros du système  $\{f_1 = \dots = f_n = 0\}$ , alors les polynômes de la RUR, calculés par rapport à cette forme linéaire, sont donnés par :

$$\begin{cases} f(t) = R(t, 1, a_1, \dots, a_n), \\ g_i(t) = \frac{\partial R}{\partial u_i}(t, 1, a_1, \dots, a_n) / \gcd(f, \frac{\partial f}{\partial t}), \quad i = 1, \dots, n. \end{cases}$$

Voir [Ren89] pour une preuve des formules ci-dessus et plus de détails.

### E.3 Complexité algorithmique et binaire d'une RUR

En utilisant les formules ci-dessus et les Théorèmes E.1 et E.2, nous pouvons estimer la taille du premier polynôme de la RUR d'un système de polynômes non-homogènes  $\{f_1, \dots, f_n\}$ .

**Théorème E.3.** Soit  $f_1, \dots, f_n$  un ensemble de polynômes non-homogènes en les variables  $x_1, \dots, x_n$ , chacun de degré  $d_i$  et de coefficients de taille binaire au plus  $\tau_i$ . Soit  $a_1 x_1 + \dots + a_n x_n$  une forme linéaire séparante des zéros de  $f_1, \dots, f_n$ , de taille binaire borné par  $\tau_s$ . Alors, le premier polynôme  $f(t)$  de la RUR a un degré en  $t$  égal à  $\prod_{i=1}^n d_i$  et son taille binaire est en  $O\left(\left(\sum_{j=1}^n \tau_j \prod_{i \neq j} d_i\right) + \left(\tau_s \prod_{i=1}^n d_i\right)\right)$ .

*Démonstration.* La preuve est immédiate en utilisant la définition de  $f$ . En effet, après homogénéisation par la variable  $x_{n+1}$ , nous obtenons de nouveaux polynômes  $F_1, \dots, F_n$ , de même degré et taille binaire que  $f_1, \dots, f_n$  auxquels nous ajoutons le polynôme  $U := u_0 t + u_1 x_1 + \dots + u_n x_n$ . Alors, d'après le Théorème E.1, le résultant de  $U, F_1, \dots, F_n$  est un polynôme en les variables  $t, u_0, \dots, u_n$  dont le degré en ces variables est  $\prod_{i=1}^n d_i$  et de taille binaire en  $O\left(\sum_{j=1}^n \tau_j \prod_{i \neq j} d_i\right)$ . En substituant les variables  $u_0, u_1, \dots, u_n$  par  $1, a_1, \dots, a_n$  dans ce polynôme, nous obtenons alors un polynôme en  $t$  de degré  $\prod_{i=1}^n d_i$  et de taille binaire en  $O\left(\left(\sum_{j=1}^n \tau_j \prod_{i \neq j} d_i\right) + \left(\tau_s \prod_{i=1}^n d_i\right)\right)$ .  $\square$

Afin d'estimer la taille des autres polynômes de la RUR, nous utilisons le résultat suivant de Mignotte [Mig89].

**Théorème E.4.** Soit  $f \in \mathbb{Z}[x]$  de degré  $d$  et de taille binaire  $\tau$ . Soit  $g \in \mathbb{Z}[x]$  un polynôme divisant  $f$  dans  $\mathbb{Z}[x]$ . Alors, la taille binaire de  $g$  est en  $O(d + \tau)$ .

**Théorème E.5.** En utilisant les notations du Théorème E.3, chaque polynôme  $g_i$  de la RUR a un degré en  $t$  borné par  $\prod_{i=1}^n d_i - 1$  et de taille binaire en  $O\left(\left(\sum_{i=1}^n \tau_j \prod_{i \neq j} d_i\right) + \left(\tau_s \prod_{i=1}^n d_i\right) + \left(\prod_{i=1}^n d_i\right)\right)$ .

*Démonstration.* Ce théorème est une conséquence directe de la définition des  $g_i$  et du Théorème E.4.  $\square$

Dans le cas où  $f_1, \dots, f_n$  ont le même degré  $d$  et taille binaire  $\tau$ , et considérant une forme linéaire séparante de même taille binaire  $\tau$ , alors les Théorèmes E.3 et E.5 donnent le résultat suivant.

**Corollaire E.1.** Les polynômes de la RUR ont le même degré en  $t$  borné par  $d^n$  et leurs coefficients ont un taille binaire borné en taille par  $O(\tau(d^n + n d^{n-1}) + d^n)$ .

## E.4 Application au système $\mathcal{E}$

Retournons maintenant au problème initial, c'est-à-dire à l'estimation de la taille de la RUR du système (E.1). Les  $\mathcal{B}_k$  sont des polynômes en les  $n$  variables  $b_0, \dots, b_{n-1}$  en les  $n$  paramètres  $d_{2k}$ . Leur degré en chacune des variables est 2. En appliquant les résultats précédents à ces polynômes, nous obtenons alors le théorème suivant.

**Théorème E.6.** Les polynômes de la RUR du système  $\{\mathcal{B}_k = 0\}_{k=0, \dots, n-1}$  ont leur degré en  $b_{n-1}$  borné par  $2^n$ . Leurs coefficients sont des polynômes en les paramètres  $d_{2k}$ ,  $k = 0, \dots, n-1$ , de degré borné par  $n 2^{n-1}$ , et de coefficients rationnels de taille binaire en  $O(2^n + n 2^{n-1})$ .

*Démonstration.* Il suffit d'appliquer les Théorèmes E.3 et E.4. □

**Remarque E.1.** Afin de réduire le coût en calcul, nous pouvons ignorer la première équation  $b_0^2 - d_0 = 0$  et considérer  $b_0$  comme un paramètre dans les autres équations. Dans ce cas, en appliquant le théorème précédent, nous obtenons que les polynômes de la RUR ont un degré en  $b_{n-1}$  borné par  $2^{n-1}$  et leurs coefficients sont des polynômes en les paramètres  $b_0, d_{2k}$ ,  $k = 1, \dots, n-1$ , de degré borné par  $(n-1) 2^{n-2}$ , et de coefficients rationnels de taille binaire en  $O((n-1) 2^{n-2})$ .

**Exemple E.1.** Soit  $n = 11$ . Alors, les polynômes de la RUR du système  $\{\mathcal{B}_k = 0\}_{k=0, \dots, 10}$  ont leur degré en  $b_{10}$  borné par  $2^{11} = 2048$ . Leurs coefficients sont des polynômes en les paramètres  $d_{2k}$ ,  $k = 0, \dots, 10$ , de degré borné par  $11 2^{10} = 11264$ .

# Annexe F

## Modélisation d'un viseur en 2 dimensions

Dans cette annexe, nous présentons une modélisation simple d'un viseur en deux dimensions, décrit par le schéma de la Figure F.1. Une telle représentation permet de comprendre quelques phénomènes physiques liés aux viseurs comme nous l'avons vu en introduction de ce mémoire. Nous proposons ici de détailler les calculs permettant d'obtenir les fonctions de transfert entre le couple moteur et la rotation de la ligne de visée. Nous utilisons ces modèles afin d'illustrer les méthodes de commande développées dans ce mémoire.

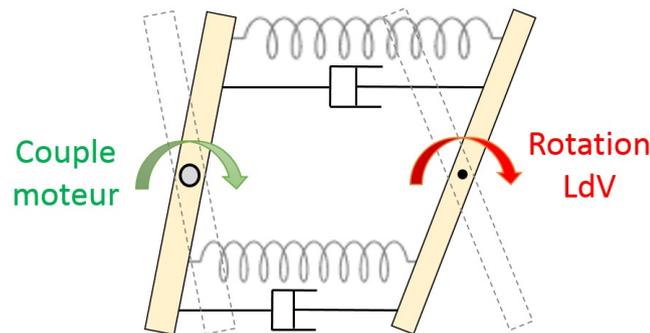


FIGURE F.1 – Modélisation d'un viseur gyrostabilisé par une double poutre

### F.1 Hypothèses et notations

Dans la Figure F.2, nous avons représenté une version plus schématique de la représentation 2D du viseur établie en Figure-F.1 afin d'introduire nos notations.

Le viseur simplifié est soumis à une vibration suivant l'axe  $\vec{x}$ . Il est composé de deux barres de longueur  $2l$ . La première barre a une masse  $m_1$ . Son centre de gravité est noté  $A$ , situé à une distance  $d_c$  du centre de symétrie  $O$  et ayant une inertie  $J_1$  en  $A$ . La position linéaire du point  $O$  est notée  $x_1$  et la position angulaire de la première barre est  $x_2$ . La deuxième barre a une masse  $m_2$ . Son centre de gravité est noté  $B$ , situé à une distance  $d_g$  du centre de symétrie  $C$  et ayant une inertie  $J_2$  en  $B$ . La position linéaire du point  $C$  est notée  $x_3$  et la position angulaire de la deuxième barre est  $x_4 := \theta$ . Les deux barres sont reliées en leurs extrémités par deux ressorts de raideur respectivement  $k_1$  et  $k_2$ .

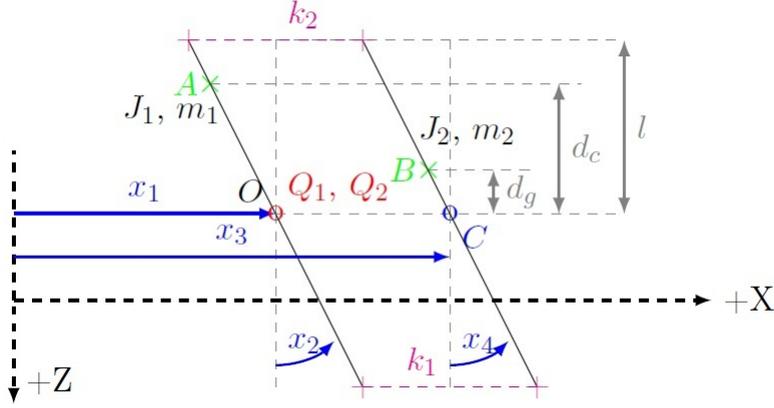


FIGURE F.2 – Schéma simplifié d'un viseur en 2D

Un moteur électrique va créer un couple  $v$  sur la première barre en  $O$ . Le couple total  $Q_2 := v = d - u$  est la différence des couples perturbateurs  $d$  et du couple moteur  $u$ . Le support vibrant va créer une force  $Q_1$  sur la première barre en  $O$ . Le capteur de position angulaire inertiel (gyroscope) ou de vitesse inertielle (gyromètre) est placé sur la deuxième barre. Ils mesurent respectivement  $x_4 := \theta$  ou  $\dot{x}_4 := \dot{\theta}$ . Un accéléromètre peut être positionné sur la deuxième barre pour mesurer l'accélération linéaire  $\ddot{x}_3$ .

Dans la suite, nous nous limiterons au cas sans balourd, c'est-à-dire à  $d_y = d_c = 0$  (voir Figure F.2). Les centres de gravité des deux poutres sont alors confondus avec les axes de rotation. Les vitesses généralisées seront notées  $\dot{x}_i := \frac{dx_i}{dt} = v_{x_i}$  ou en variable de Laplace, par  $s \hat{x}_i$ . Les accélérations généralisées seront notées  $\ddot{x}_i := \frac{d^2 x_i}{dt^2} = a_{x_i}$  ou, en variable de Laplace par  $s^2 \hat{x}_i$ . La première barre est entraînée en vibration linéaire  $\ddot{x}_1$ , ce qui va induire une accélération linéaire  $\ddot{x}_3$  sur la deuxième. Comme, par hypothèse  $k_1 \neq k_2$ , cette vibration linéaire va aussi produire une rotation  $x_4 := \theta$  de la deuxième barre. Les mouvements linéaires vont créer des mouvements angulaires. Quand un couple  $Q_2 := v$  est appliqué à la première barre, celle-ci se met en rotation. Par contre, l'inertie de la deuxième barre tend à immobiliser cette dernière. Il en résulte une déformation au niveau des ressorts.

## F.2 Équations de la mécanique

L'objectif de ce paragraphe est de trouver la relation qui relie la rotation  $x_4$  (visible par un gyromètre) au couple  $Q_2$ . Pour simplifier les calculs, dans un premier temps, nous allons négliger l'amortissement des ressorts  $f_1 = f_2 = 0$ . À partir de l'énergie cinétique

$$2T = m_1 \dot{x}_1^2 + J_1 \dot{x}_2^2 + m_2 \dot{x}_3^2 + J_2 \dot{x}_4^2,$$

et de l'énergie potentielle

$$2V = k_1 (x_3 + l \sin(x_4) - x_1 - l \sin(x_2))^2 + k_2 (x_3 - l \sin(x_4) - x_1 + l \sin(x_2))^2,$$

nous pouvons définir le *Lagrangien* comme la différence entre l'énergie cinétique et l'énergie potentielle :

$$L := T - V. \quad (\text{F.1})$$

Les forces généralisées  $Q_i$  vont créer un *travail virtuel*  $W$ . Dans notre cas, les deux seules forces généralisées sont d'une part le couple total au niveau de la motorisation  $Q_2$  et d'autre

part la force de réaction du porteur  $Q_1$ . Elles interviennent respectivement sur les déplacements généralisés  $x_2$  et  $x_1$ . Nous obtenons alors :

$$\begin{cases} \frac{\partial W}{\partial x_1} = \frac{\partial}{\partial x_1} (Q_1 x_1) = Q_1, \\ \frac{\partial W}{\partial x_2} = \frac{\partial}{\partial x_2} (Q_2 x_2) = Q_2, \\ Q_i = 0, \quad i \neq 1, 2. \end{cases}$$

Pour chaque degré de liberté  $x_i$  et vitesse  $\dot{x}_i$ , les 4 équations d'*Euler-Lagrange* sont :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} = Q_i, \quad i = 1, \dots, 4. \quad (\text{F.2})$$

Après linéarisation (en prenant de petits angles et en négligeant les produits de vitesse), nous obtenons le système d'équations ci-dessous (en remplaçant la dérivation par la variable  $s$  de Laplace) :

$$\begin{cases} (k_1 + k_2 + m_1 s^2) \hat{x}_1 + (k_1 - k_2) l \hat{x}_2 - (k_1 + k_2) \hat{x}_3 - (k_1 - k_2) l \hat{x}_4 = \hat{Q}_1, \\ (k_1 - k_2) l \hat{x}_1 + (J_1 s^2 + (k_1 + k_2) l^2) \hat{x}_2 - (k_1 - k_2) l \hat{x}_3 - (k_1 + k_2) l^2 \hat{x}_4 = \hat{Q}_2, \\ -(k_1 + k_2) \hat{x}_1 - (k_1 - k_2) l \hat{x}_2 + (m_2 s^2 + k_1 + k_2) \hat{x}_3 + (k_1 - k_2) l \hat{x}_4 = 0, \\ -(k_1 - k_2) l \hat{x}_1 - (k_1 + k_2) l^2 \hat{x}_2 + (k_1 - k_2) l \hat{x}_3 + (J_2 s^2 + (k_1 + k_2) l^2) \hat{x}_4 = 0. \end{cases}$$

Pour introduire l'effet des amortissements visqueux ( $f_1, f_2$ ) des deux ressorts, opérons le changement de variable suivant :

$$k_1 \rightarrow k_1 + f_1 s, \quad k_2 \rightarrow k_2 + f_2 s.$$

Finalement le système mécanique des deux poutres est modélisé par 4 équations différentielles ordinaires du second ordre suivantes :

$$\begin{cases} (m_1 s^2 + (f_1 + f_2) s + k_1 + k_2) \hat{x}_1 + ((f_1 - f_2) l s + (k_1 - k_2) l) \hat{x}_2 \\ + ((-f_1 - f_2) s - k_1 - k_2) \hat{x}_3 + ((-f_1 + f_2) l s + (-k_1 + k_2) l) \hat{x}_4 = \hat{Q}_1, \\ ((f_1 - f_2) l s + (k_1 - k_2) l) \hat{x}_1 + (J_1 s^2 + (f_1 + f_2) l^2 s + (k_1 + k_2) l^2) \hat{x}_2 \\ + ((-f_1 + f_2) l s + (-k_1 + k_2) l) \hat{x}_3 + ((-f_1 - f_2) l^2 s + (-k_1 - k_2) l^2) \hat{x}_4 = \hat{Q}_2, \\ ((-f_1 - f_2) s - k_1 - k_2) \hat{x}_1 + ((-f_1 + f_2) l s + (-k_1 + k_2) l) \hat{x}_2 \\ + (m_2 s^2 + (f_1 + f_2) s + k_1 + k_2) \hat{x}_3 + ((f_1 - f_2) l s + (k_1 - k_2) l) \hat{x}_4 = 0, \\ ((-f_1 + f_2) l s + (-k_1 + k_2) l) \hat{x}_1 + ((-f_1 - f_2) l^2 s + (-k_1 - k_2) l^2) \hat{x}_2 \\ + ((f_1 - f_2) l s + (k_1 - k_2) l) \hat{x}_3 + (J_2 s^2 + (f_1 + f_2) l^2 s + (k_1 + k_2) l^2) \hat{x}_4 = 0. \end{cases}$$

Comme le mouvement de  $x_1$  est imposé par l'extérieur du système et que nous ne sommes pas intéressés par  $Q_1$ , il est possible d'exprimer  $x_2, x_3, x_4$  en fonction de  $Q_2$  et  $x_1$ . Maintenant, si l'on pose

$$\begin{aligned} \Delta f &:= f_1 - f_2, & \Delta k &:= k_1 - k_2, & \Sigma f &:= f_1 + f_2, & \Sigma k &:= k_1 + k_2, \\ a &:= \Sigma f s + \Sigma k, & b &:= (\Delta f s + \Delta k) l, \end{aligned}$$

nous obtenons la matrice de masse suivante

$$M = \begin{pmatrix} J_1 s^2 + a l^2 & -b & -a l^2 \\ -b & m_2 s^2 + a & b \\ -a l^2 & b & J_2 s^2 + a l^2 \end{pmatrix},$$

qu'il est possible d'inverser en utilisant son déterminant, défini par :

$$\det(M) = m_2 J_1 J_2 s^6 + ((J_1 + J_2) l^2 m_2 + J_1 J_2) a s^4 + ((J_1 + J_2) l^2 a^2 - (J_1 + J_2) b^2) s^2 .$$

**Proposition F.1.** Le système linéaire correspondant à la modélisation du viseur 2D précédent, dont les variables de sortie sont  $x_2$  (la rotation du cadre),  $x_3$  et  $x_4$  (respectivement la translation et la rotation du cœur), et les entrées  $Q_2$  (le couple total) et  $x_1$  (le déplacement du support), est défini par

$$\begin{pmatrix} \widehat{x}_2 \\ \widehat{x}_3 \\ \widehat{x}_4 \end{pmatrix} = M^{-1} \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \widehat{Q}_2 + \begin{pmatrix} -b \\ a \\ b \end{pmatrix} \widehat{x}_1 \right), \quad M^{-1} = \frac{N^\dagger}{(J_1 + J_2) s^2 D^\dagger}, \quad (\text{F.3})$$

avec

$$N^\dagger := \begin{pmatrix} J_2 m_2 s^4 + (J_2 + m_2 l^2) a s^2 + \alpha & J_2 b s^2 & m_2 a l^2 s^2 + \alpha \\ J_2 b s^2 & J_1 J_2 s^4 + (J_1 + J_2) a l^2 s^2 & -J_1 b s^2 \\ m_2 a l^2 s^2 + \alpha & -J_1 b s^2 & J_1 m_2 s^4 + (J_1 + m_2 l^2) a s^2 + \alpha \end{pmatrix},$$

$$D^\dagger := \alpha + \left( \frac{J_1 J_2}{J_1 + J_2} + m_2 l^2 \right) a s^2 + \frac{J_1 J_2}{J_1 + J_2} m_2 s^4,$$

et :

$$\begin{aligned} \Delta f &:= f_1 - f_2, & \Delta k &:= k_1 - k_2, & \Sigma f &:= f_1 + f_2, & \Sigma k &:= k_1 + k_2, \\ a &:= \Sigma f s + \Sigma k, & b &:= (\Delta f s + \Delta k) l, \\ \alpha &:= a^2 l^2 - b^2 = 4 k_1 k_2 l^2 + 4 (f_1 k_2 + f_2 k_1) l^2 s + 4 f_1 f_2 l^2 s^2. \end{aligned}$$

## F.3 Flexibilité

Nous allons nous intéresser maintenant à la fonction de transfert de la mesure du gyroscope  $x_4$  en fonction du couple total  $Q_2$ . Cette fonction de transfert  $P = \frac{\widehat{x}_4}{\widehat{Q}_2} \Big|_{\widehat{x}_1=0}$  s'appelle la *flexibilité*.

### F.3.1 Flexibilité exacte

Pour calculer la flexibilité, il faut poser  $\widehat{x}_1 = 0$  dans l'équation (F.3) et ne prendre en compte que les termes  $N_{3,1}^\dagger$  et  $D^\dagger$ , puis introduire la nouvelle variable :

$$\beta := \frac{\alpha}{4 k_1 k_2 l^2} = 1 + \left( \frac{f_1}{k_1} + \frac{f_2}{k_2} \right) s + \frac{f_1 f_2}{k_1 k_2} s^2 .$$

Nous avons alors que l'angle du gyroscope est défini en fonction du couple total par

$$P := \frac{\widehat{x}_4}{\widehat{Q}_2} = \frac{N}{(J_1 + J_2) s^2 D},$$

où le numérateur vaut

$$N := \frac{N_{3,1}^\dagger}{4 k_1 k_2 l^2} = 1 + \left( \frac{f_1}{k_1} + \frac{f_2}{k_2} \right) s + \left( \frac{f_1 f_2}{k_1 k_2} + m_2 \frac{k_1 + k_2}{4 k_1 k_2} \right) s^2 + m_2 \frac{f_1 + f_2}{4 k_1 k_2} s^3,$$

et le dénominateur est le système du 6<sup>e</sup> ordre défini par :

$$\begin{aligned} D &:= \frac{D^\dagger}{4 k_1 k_2 l^2} \\ &= 1 + \left( \frac{f_1}{k_1} + \frac{f_2}{k_2} \right) s + \left( \frac{f_1 f_2}{k_1 k_2} + \left( \frac{J_1 J_2}{J_1 + J_2} + m_2 l^2 \right) \frac{k_1 + k_2}{4 k_1 k_2 l^2} \right) s^2 \\ &\quad + \left( \frac{J_1 J_2}{J_1 + J_2} + m_2 l^2 \right) \frac{f_1 + f_2}{4 k_1 k_2 l^2} s^3 + \frac{J_1 J_2}{J_1 + J_2} \frac{m_2}{4 k_1 k_2 l^2} s^4. \end{aligned}$$

**Corollaire F.1.** Posons les nouvelles inertie et raideur suivantes :

$$J_e := \frac{J_1 J_2}{J_1 + J_2}, \quad k_e := 4 \frac{k_1 k_2}{k_1 + k_2} l^2.$$

Alors, nous avons la flexibilité avec 2 modes mécaniques définie par la fonction de transfert suivante :

$$\begin{aligned} P &:= \frac{\widehat{x}_4}{\widehat{Q}_2} \Big|_{\widehat{x}_1=0} = \frac{N}{(J_1 + J_2) s^2 D}, \\ N &:= 1 + \left( \frac{f_1}{k_1} + \frac{f_2}{k_2} \right) s + \left( \frac{f_1 f_2}{k_1 k_2} + \frac{m_2 l^2}{k_e} \right) s^2 + \frac{m_2 l^2}{k_e} \frac{f_1 + f_2}{k_1 + k_2} s^3, \\ D &:= 1 + \left( \frac{f_1}{k_1} + \frac{f_2}{k_2} \right) s + \left( \frac{f_1 f_2}{k_1 k_2} + \frac{J_e + m_2 l^2}{k_e} \right) s^2 + \frac{J_e + m_2 l^2}{k_e} \frac{f_1 + f_2}{k_1 + k_2} s^3 + \frac{J_e m_2}{k_e (k_1 + k_2)} s^4. \end{aligned} \tag{F.4}$$

La flexibilité fait apparaître une *inertie pure* en série avec un filtre passe-bas du quatrième ordre.

### F.3.2 Cas idéal de flexibilité pour deux ressorts égaux

Même dans le cas idéal où les deux ressorts sont égaux ( $k_1 = k_2$ ), l'équation (F.4) fait quand même apparaître un mode de rotation qui va "gêner" l'asservissement. A partir du résultat précédent (Corollaire F.1) dans le cas  $k_e = 2 k l^2$ , nous remarquons qu'il est possible de factoriser le numérateur

$$N = 1 + 2 \frac{f}{k} s + \left( \frac{f^2}{k^2} + \frac{m_2}{2k} \right) s^2 + \frac{f m_2}{2 k^2} s^3 = \left( 1 + \frac{f}{k} s \right) \left( 1 + \frac{f}{k} s + \frac{m_2}{2k} s^2 \right),$$

et le dénominateur :

$$\begin{aligned} D &= 1 + 2 \frac{f}{k} s + \left( \frac{m_2}{2k} + \frac{f^2}{k^2} + \frac{J_e}{k_e} \right) s^2 + \left( \frac{f m_2}{k} + \frac{f J_e}{k k_e} \right) s^3 + \frac{J_e m_2}{k_e} \frac{1}{2k} s^4 \\ &= \left( 1 + \frac{f}{k} s + \frac{J_e}{2k l^2} s^2 \right) \left( 1 + \frac{f}{k} s + \frac{m_2}{2k} s^2 \right). \end{aligned}$$

La masse  $m_2$  n'intervient alors plus et la modélisation se simplifie encore.

**Corollaire F.2.** Dans le cas d'un système symétrique, c'est-à-dire avec  $k_1 = k_2 = k$  et  $f_1 = f_2 = f$ , l'inertie équivalente est  $J_e := \frac{J_1 J_2}{J_1 + J_2}$  et la nouvelle raideur équivalente est  $k_e = 2k l^2$ . On définit la pulsation  $p := \sqrt{\frac{k_e}{J_e}}$ , l'amortissement  $\epsilon := \frac{f p}{2k}$  et l'inverse de l'inertie

totale  $n_0 := \frac{1}{J_1+J_2}$ . La flexibilité se simplifie alors en

$$\begin{cases} P := \frac{\hat{x}_4}{\widehat{Q}_2} \Big|_{x_1=0} = \frac{1}{(J_1 + J_2) s^2} \frac{1 + \frac{f}{k} s}{1 + \frac{f}{k} s + \frac{J_e}{2kl^2} s^2} = \frac{n_0 (p^2 + 2\epsilon p s)}{s^2 (p^2 + 2\epsilon p s + s^2)}, \\ G := \frac{n_0 (p^2 + 2\epsilon p s)}{s (p^2 + 2\epsilon p s + s^2)}, \end{cases} \quad (\text{F.5})$$

qui fait apparaître une *inertie pure en série avec un filtre passe-bas du second ordre*. Ce filtre passe-bas correspond au mode mécanique de plus basse fréquence du viseur. Ce mode est très limitant en terme de performance de l'asservissement.

### F.3.3 Cas de deux ressorts égaux sans amortissement

Nous reprenons l'équation (F.5) de la modélisation en transfert  $P$  entre la position inertielle  $x_4 := \theta$  et le couple total  $Q_2 := v$  (qui est la somme du couple moteur  $u$  et des couples perturbateurs  $d$ ). Nous faisons le choix de considérer le mode de fréquence  $p$  comme non amorti (Corollaire F.2) en imposant  $\epsilon := f/k = 0$ , c'est-à-dire  $f = 0$ . Cela donne une nouvelle modélisation  $\widehat{P}$  du système réel  $P$ . Cette hypothèse permet de se placer dans *le pire des cas*, dans le sens où si nous pouvons concevoir un correcteur stabilisant un tel système, alors ce même correcteur stabilisera le même système avec un amortissement non nul.

**Proposition F.2.** Pour un système décrit par la modélisation double poutre dans le cas sans balourd et symétrique, c'est-à-dire où les deux ressorts sont de même raideur  $k$  et sans amortissement  $f = 0$  alors l'inertie équivalente est  $J_e := \frac{J_1 J_2}{J_1 + J_2}$  et la nouvelle raideur équivalente devient  $k_e = 2kl^2$ . On définit alors la pulsation  $p := \sqrt{\frac{k_e}{J_e}}$  et l'inverse de l'inertie totale  $n_0 := \frac{1}{J_1 + J_2}$ . La flexibilité entre la position inertielle  $x_4 := \theta$  et le couple total  $Q_2 := v = -u + d$ , qui est la somme du couple moteur  $u$  et des couples perturbateurs  $d$ , se modélise alors par :

$$\begin{cases} \widehat{P} := \frac{\hat{x}_4}{\widehat{Q}_2} = \frac{1}{(J_1 + J_2) s^2} \frac{1}{1 + \frac{s^2}{p^2}} = \frac{n_0 p^2}{s^2 (p^2 + s^2)}, \\ \widehat{G} := \widehat{P} s = \frac{n_0 p^2}{s (p^2 + s^2)}. \end{cases} \quad (\text{F.6})$$

Ce système fait apparaître une *inertie pure en série avec un filtre-passe bas du second ordre sans amortissement correspondant au premier mode limitant le bouclage du viseur*.

# Annexe G

## Identification de fonction de transfert

Considérons la stabilisation d'une ligne de visée d'un viseur "miror stabilized", c'est-à-dire dont la ligne de visée est stabilisée à l'aide de la motorisation de miroirs, telle que décrite sur le schéma de la Figure G.1. La problématique majeure de ce système vient des courroies mises en place pour la stabilisation. Les courroies sont des structures mécaniques très flexibles qui entrent en résonance à certaines fréquences. De plus, ces fréquences de résonance sont susceptibles de varier d'un viseur à l'autre. Ceci motive les perspectives en commande adaptative développées en conclusion de cette thèse.

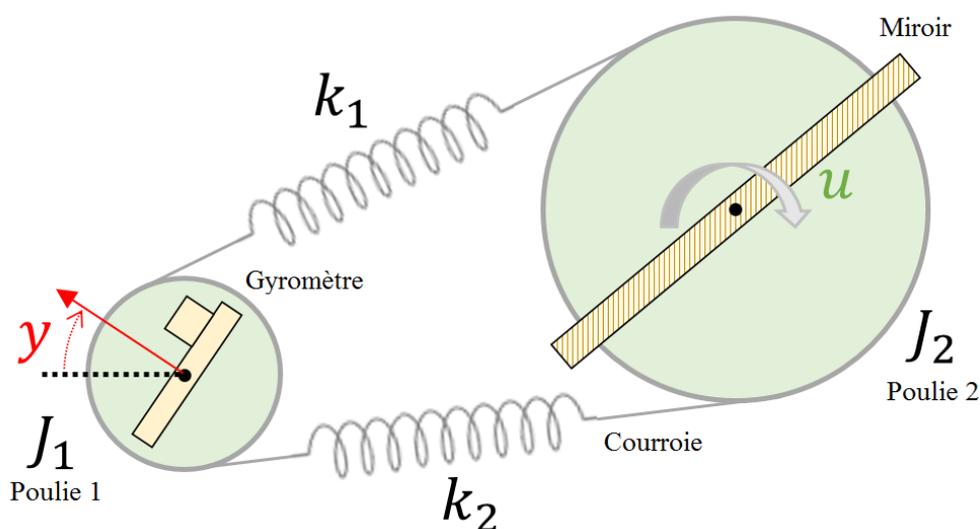


FIGURE G.1 – Système de stabilisation *Bloc Gyro Miroir (BGM)*

L'objectif de cette annexe est de montrer comment identifier une fonction de transfert à partir de données expérimentales mesurées sur un prototype réel. Dans un souci de confidentialité, les données sont normalisées.

### G.1 Mesures hors ligne à disposition

Nous disposons de deux jeux de mesures :

- Des mesures de fonctions de transfert entre la tension moteur  $u$  et la vitesse angulaire de la Ligne de Visée  $y$  (voir Figure G.2).

- Des mesures de fonction de transfert entre le courant moteur  $i$  et la vitesse angulaire de la Ligne de Visée  $y$  (voir Figure G.3).

Ces mesures mettent en évidence la présence de deux modes mécaniques très prononcés.

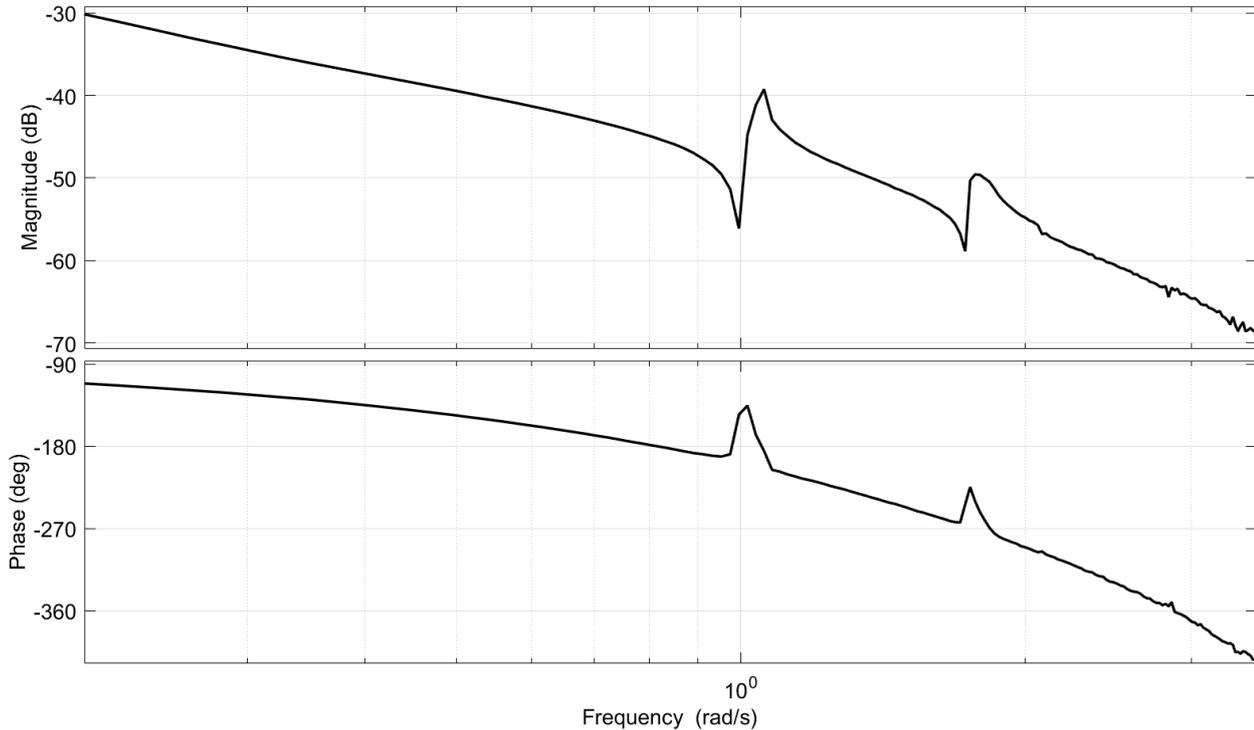


FIGURE G.2 – Réponse fréquentielle de  $y/u = G_m P H$

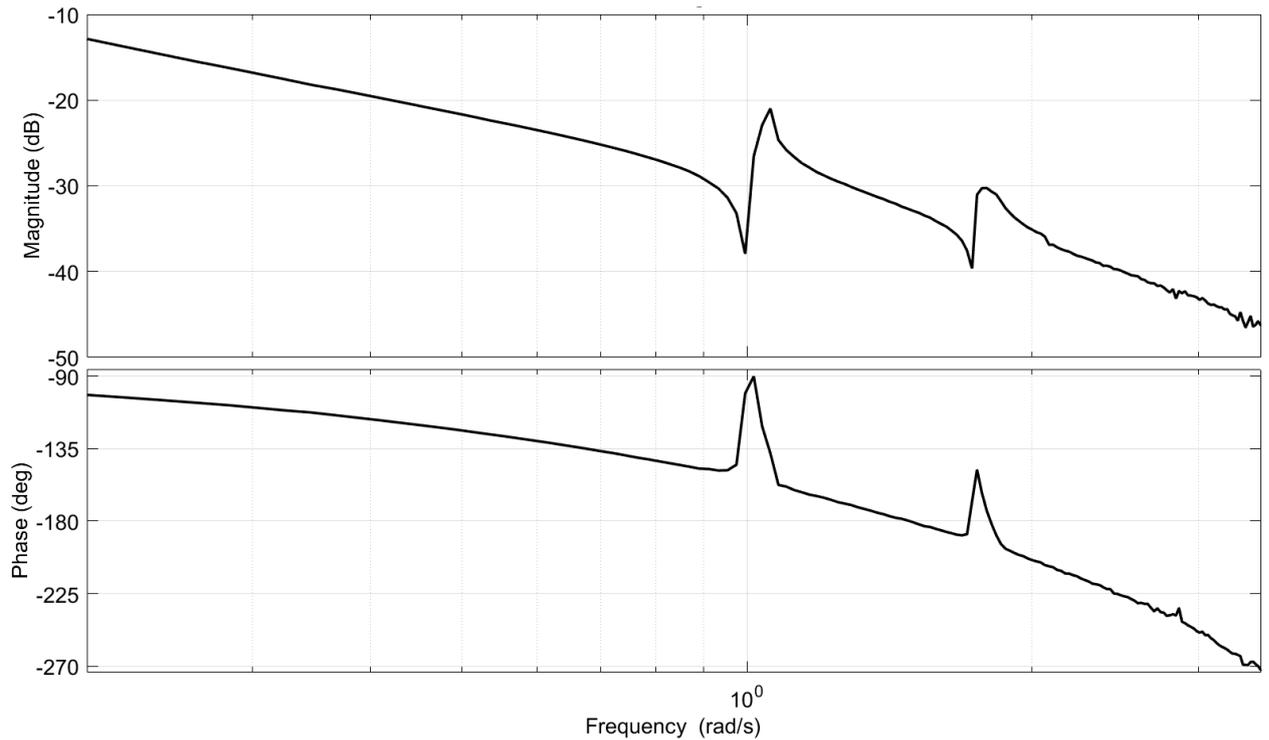
## G.2 Identification du système

Bien que de nombreux algorithmes soient à disposition de l'ingénieur automatique, l'identification de la fonction de transfert  $G$  entre commande moteur  $u$  et vitesse angulaire mesurée  $y$  n'est pas chose aisée. Ces algorithmes permettent en général une bonne identification mais proposent généralement des fonctions de transfert d'ordre élevé afin d'identifier correctement  $G$ . Nous pouvons cependant nous appuyer sur notre connaissance de la physique de notre système ainsi que de nos deux mesures afin de simplifier ces modèles identifiés. Dans le cas des viseurs, le système  $G$  peut se décomposer en trois sous-systèmes : une partie moteur, une partie capteur et une partie mécanique (voir Section 1.3). Identifier la fonction de transfert  $G$  revient à identifier la fonction de transfert de chaque sous-système.

Sachant qu'en basse fréquence, le système se ramène à une simple inertie mécanique, l'ajustement "à la main" d'un gain  $G_0$  permettant de "faire coïncider" la mesure en tension avec  $\frac{G_0}{s}$  permet de déduire grossièrement le gain statique du modèle  $G$ . Nous retenons :

$$G_0 := 6.65 \cdot 10^{-3}.$$

Aussi, l'utilisation combinée des deux mesures en courant et tension (Figures G.3 et G.2) nous permet de déduire un modèle pour la partie moteur d'ordre 1 ou 2. Connaissant cela,

FIGURE G.3 – Réponse fréquentielle de  $y/i = K_m P H$ 

il reste à déterminer la partie capteur. L'ajustement du gain permet d'identifier l'ordre de la fonction de transfert. Nous pouvons ensuite recalibrer la phase avec un retard pur (ou une approximation de celui-ci).

Cependant, si l'on cherche à obtenir des modélisations d'ordre le plus faible possible, il est alors préférable d'identifier les parties moteur et capteur comme une seule et unique fonction de transfert. En effet, la démarche précédente, bien que respectant la connaissance de la physique du système, a pour inconvénient de donner des fonctions de transfert d'ordre un peu plus élevé que ce que l'on pourrait souhaiter.

### G.2.1 Une première modélisation simple avec retard

Nous considérons ici une modélisation simple de  $G$ , précise jusqu'à la fréquence du premier mode. Le premier mode mécanique est le plus important en pratique car c'est celui qui conditionne la bande passante du système. On peut alors modéliser  $G$  par une inertie, un mode mécanique amorti et un retard. Notons aussi que l'on considère des amortissements identiques pour la résonance et l'anti-résonance du mode mécanique. Cette simplification a son importance dans le Chapitre 5 de ce mémoire, et ne dégrade que très légèrement la modélisation. Nous notons cette fonction de transfert  $G_{3,r}$

$$G_{3,r} := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) e^{-T_{3,r} s},$$

$$\omega_0 = 1, \quad \omega_1 = 1.03, \quad \xi = 1.08 \cdot 10^{-2}, \quad T_{3,r} = 1.82,$$

représentée en Figure G.4. Cette modélisation revient à identifier le système par un système de deux poutres reliées par des ressorts comme nous le présentions en Annexe F auquel nous avons ajouté un retard pur.

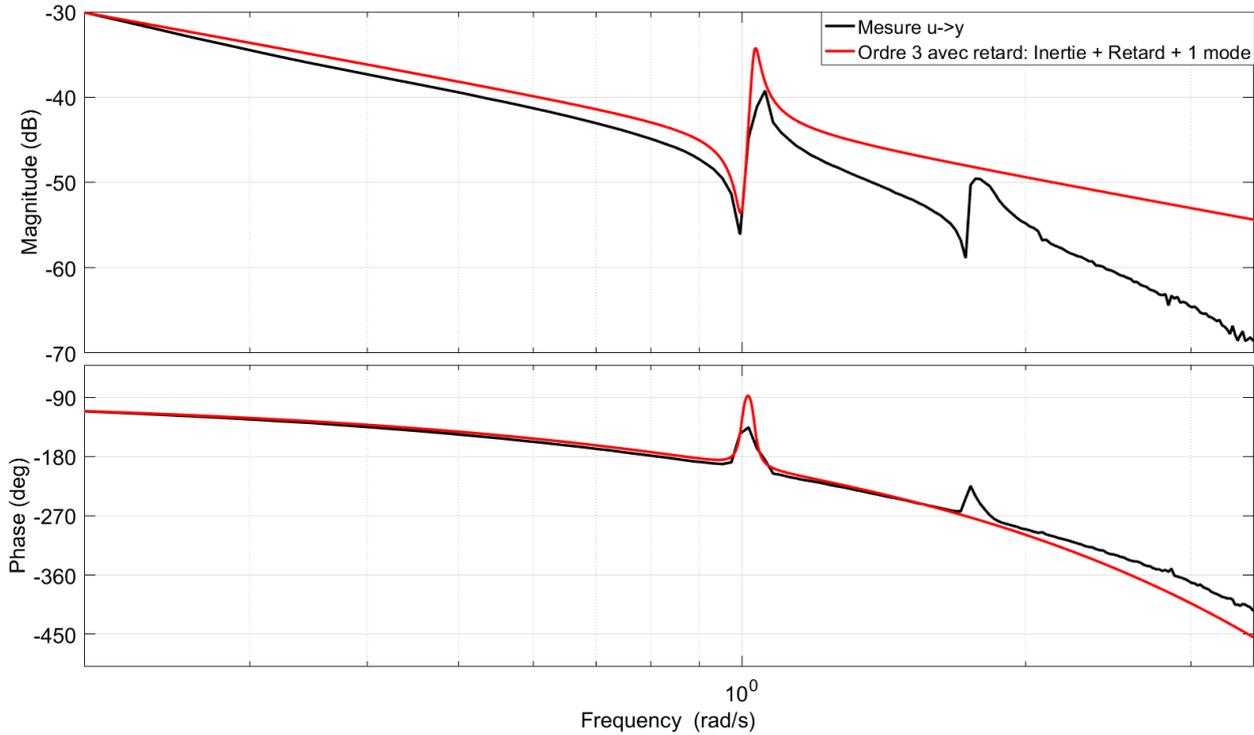


FIGURE G.4 – Modélisation avec une inertie, un mode mécanique et un retard

**Remarque G.1.** Dans un souci de compacité, nous noterons dans la suite :

$$H_{m_1} := \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2}.$$

**Remarque G.2.** La valeur de  $T_{3,r}$  se déduit rapidement à la main, en ajustant la phase. Les valeurs de  $H_{m,1}$  peuvent s'obtenir à l'aide d'algorithmes d'identification tels que `fitfrd` ou `invfreq` de MATLAB, ou à la main par essai-erreur.

## G.2.2 Une modélisation d'ordre 4 sans retard

La modélisation précédente peut encore se simplifier, en approximant le retard par une approximation de Padé d'ordre 1. Une telle approximation du retard étant peu précise en haute fréquence, nous révisons la valeur du retard et obtenons la fonction de transfert  $G_4$

$$G_4 := \frac{G_0}{s} H_{m_1} \left( \frac{1 - \frac{T_4}{2}s}{1 + \frac{T_4}{2}s} \right)$$

$$G_0 = 6.65 \cdot 10^{-3}, \quad \omega_0 = 1, \quad \omega_1 = 1.03, \quad \xi = 1.08 \cdot 10^{-2}, \quad T_4 = 2.36,$$

représentée en Figure G.5.

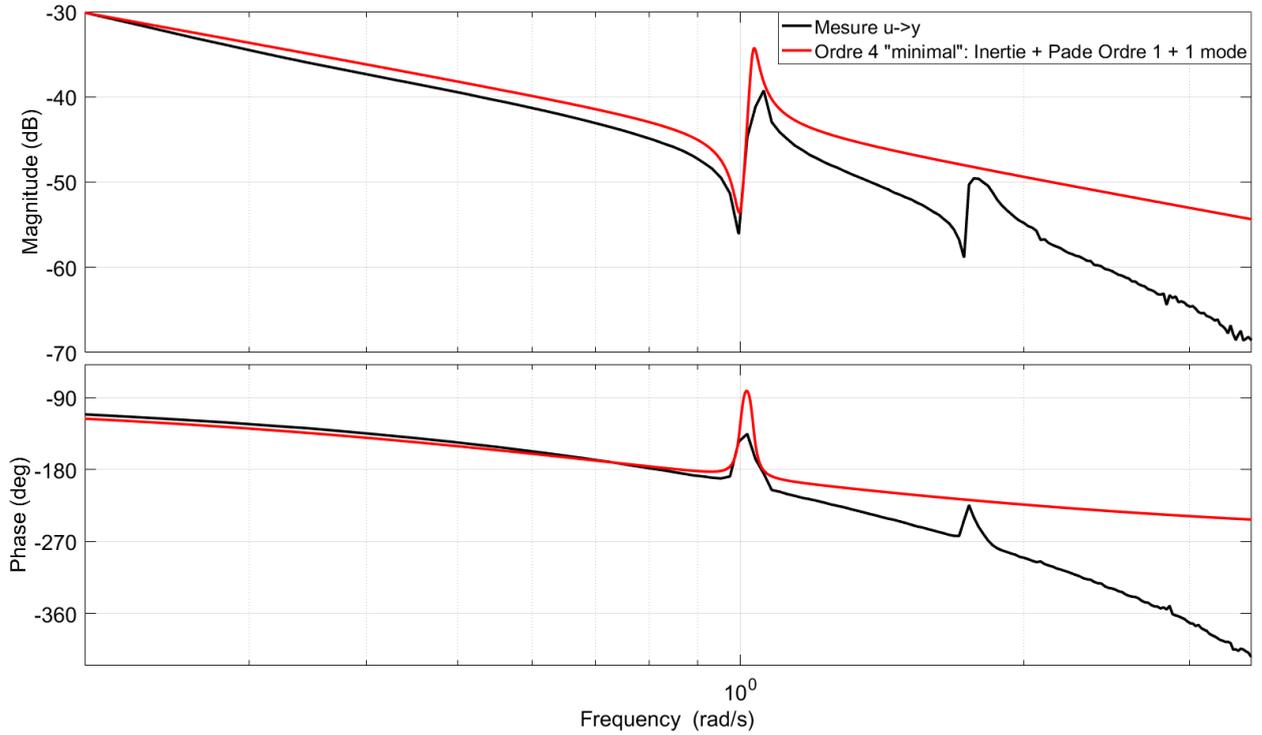


FIGURE G.5 – Modélisation avec une inertie, un mode mécanique et une approximation de Padé d’ordre 1

### G.2.3 Une modélisation d’ordre 5 plus fine

Les deux modélisations précédentes sont correctes jusqu’à la fréquence du 1<sup>er</sup> mode. On peut rajouter un effet passe-bas afin de mieux coller en gain et en phase après la fréquence du 1<sup>er</sup> mode. On choisit un passe-bas de la forme suivante :

$$F_1 := \frac{1}{1 + s/1.14}.$$

Nous obtenons alors les fonctions de transferts  $G_{4,r}$  et  $G_5$  définies par

$$\begin{cases} G_{4,r} := \frac{G_0}{s} H_{m_1} F_1 e^{-T_{4,r} s}, & \text{où } T_{4,r} = 1.21, \\ G_5 := \frac{G_0}{s} H_{m_1} F_1 \left( \frac{1 - \frac{T_5}{2} s}{1 + \frac{T_5}{2} s} \right), & \text{où } T_5 = 1.33, \end{cases}$$

et représentées en Figure G.6. La modélisation  $G_{4,r}$  sert d’exemple au Chapitre 5.

**Remarque G.3.** Les valeurs du filtre passe-bas  $F_1$  peuvent se déterminer “à la main” en réglant d’abord la fréquence de coupure du filtre, puis en réajustant le gain du filtre. Les valeurs des retards  $T_{4,r}$  et  $T_5$  se déterminent simplement en ajustant la phase du modèle total.

### G.2.4 Modélisations plus fines

Il est toujours possible de trouver des modèles plus précis mais mathématiquement plus complexes. Les possibilités de modélisation sont nombreuses. Nous retiendrons simplement

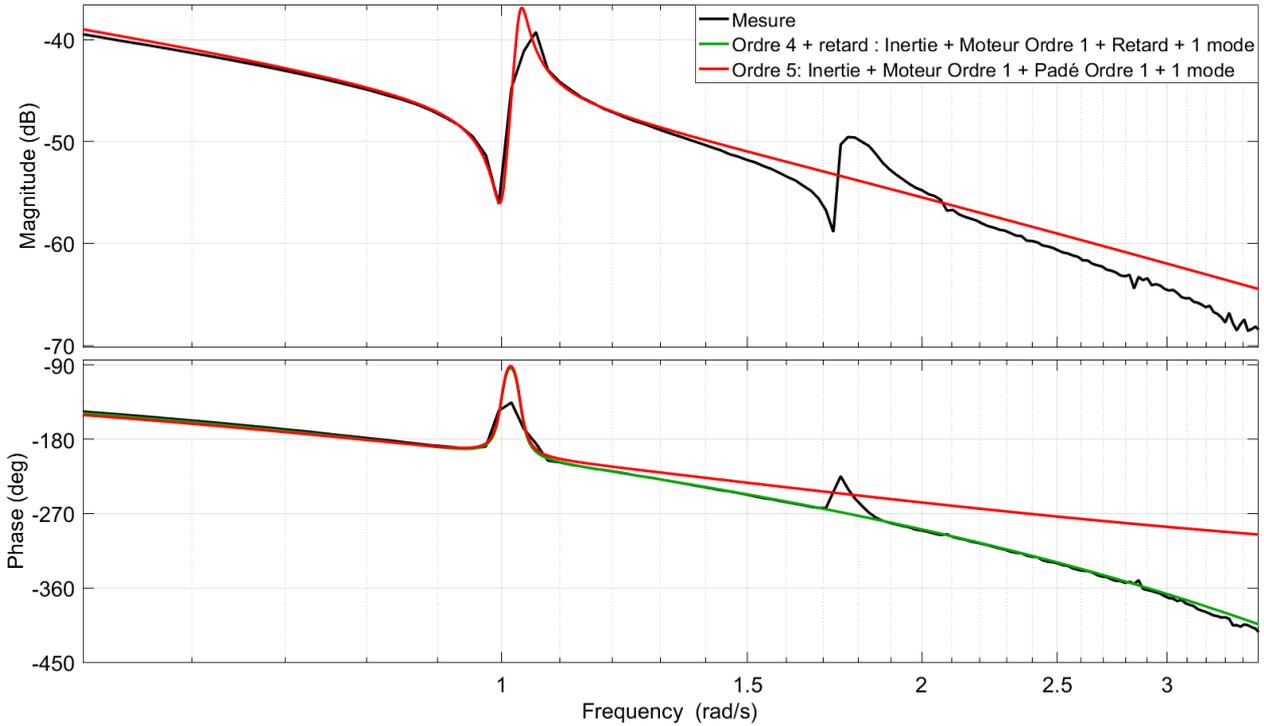


FIGURE G.6 – Modélisation avec une inertie, un mode mécanique, un passe-bas et un retard pur ou défini par une approximation de Padé d'ordre 1

deux modélisations faisant intervenir un filtre passe-bas  $F_2$  du second ordre

$$F_2 := \frac{3.0429}{3.0430 + 3.3260 s + s^2},$$

et un deuxième mode mécanique

$$H_{m_2} := \frac{s^2 + 2 \xi_2 \omega_2 s + \omega_2^2}{s^2 + 2 \xi_3 \omega_3 s + \omega_3^2},$$

$$\omega_2 = 1.74, \quad \xi_2 = 5.74 \cdot 10^{-3}, \quad \omega_3 = 1.75, \quad \xi_3 = 3.80 \cdot 10^{-3},$$

pour obtenir les fonctions de transfert suivantes :

$$\begin{cases} G_{7,r} := \frac{G_0}{s} H_{m_1} H_{m_2} F_2 e^{-T_{7,r} s}, & \text{où } T_{7,r} = 9.21 \cdot 10^{-1}, \\ G_9 := \frac{G_0}{s} H_{m_1} H_{m_2} F_2 \left( \frac{1 - \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}}{1 + \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}} \right), & \text{où } T_9 = 9.21 \cdot 10^{-1}. \end{cases}$$

Notamment, la modélisation  $G_9$  est reprise dans la Section 2.6 du Chapitre 2.

### G.3 Récapitulatif des différentes modélisations

Nous avons vu que nous pouvions choisir entre un certain nombre de modélisations. Celles-ci sont récapitulées dans le Tableau G.3 et représentées dans la Figure G.8. Plus celles-ci seront fines, plus la conception de correcteurs sera délicate. Ces modélisations nous permettront de tester les différents résultats de ce mémoire de thèse.

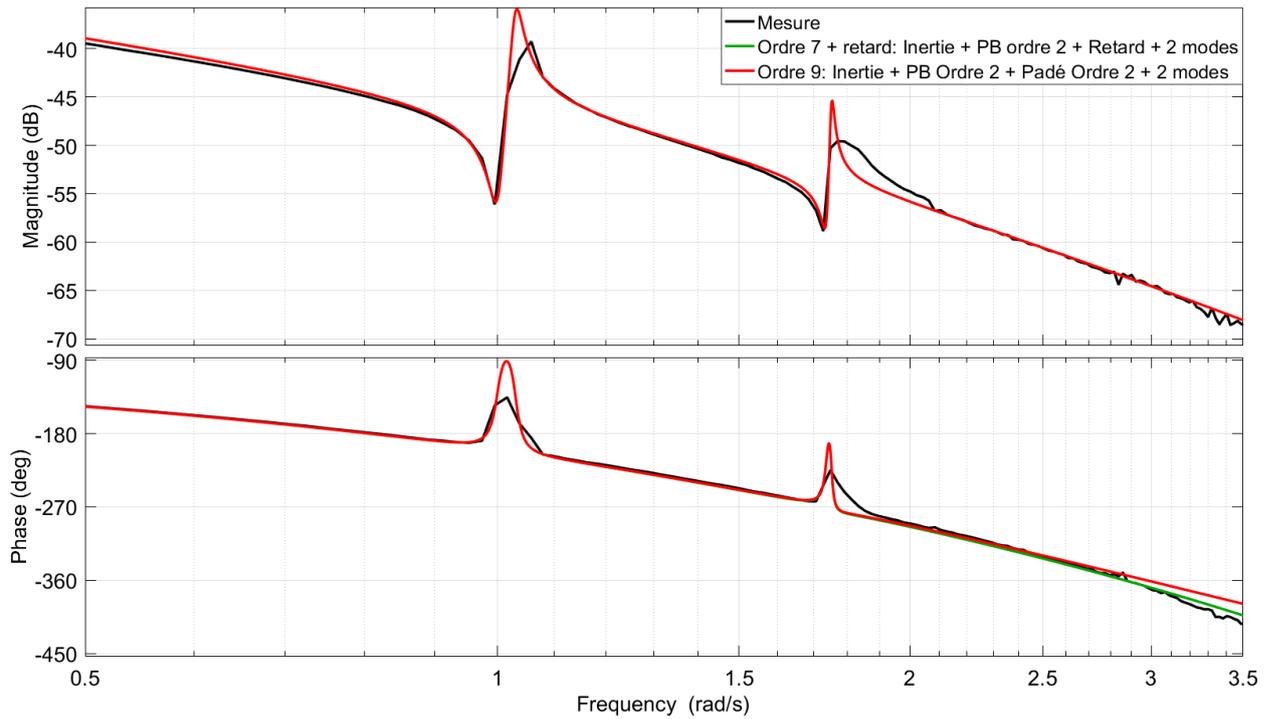


FIGURE G.7 – Modélisation avec une inertie, un mode mécanique, un passe-bas et un retard pur ou défini par une approximation de Padé d'ordre 1

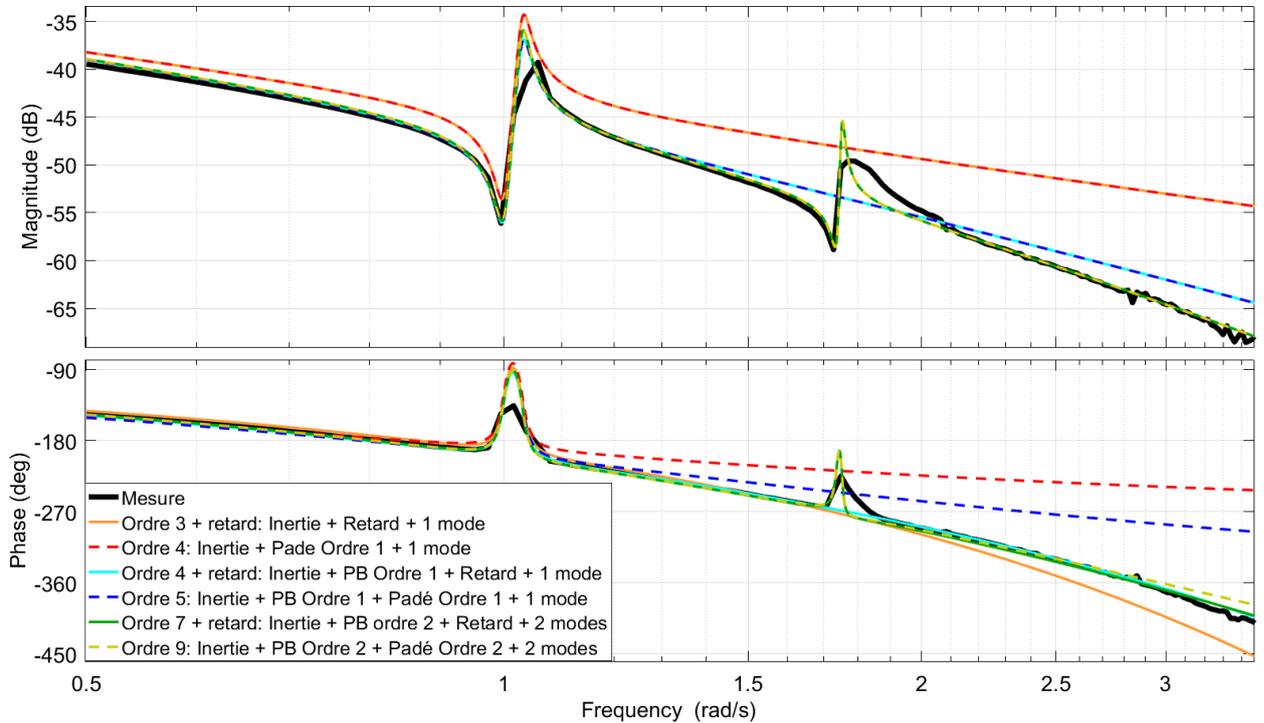


FIGURE G.8 – Récapitulatif des différentes modélisations

Modélisation	paramètres physiques normalisés
$G_{3,r} := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) e^{-T_{3,r} s}$	$G_0 = 6.65 \cdot 10^{-3}$ $T_{3,r} = 1.82$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$
$G_4 := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{1 - \frac{T_4}{2}s}{1 + \frac{T_4}{2}s} \right)$	$G_0 = 6.65 \cdot 10^{-3}$ $T_4 = 2.36$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$
$G_{4,r} := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{1}{1 + \tau_1 s} \right) e^{-T_{4,r} s}$	$G_0 = 6.65 \cdot 10^{-3}$ $T_{4,r} = 1.21$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$ $\tau_1 = 1/1.14$
$G_5 := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{1}{1 + \tau_1 s} \right) \left( \frac{1 - \frac{T_5}{2}s}{1 + \frac{T_5}{2}s} \right)$	$G_0 = 6.65 \cdot 10^{-3}$ $T_5 = 1.33$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$ $\tau_1 = 1/1.14$
$G_{7,r} := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{s^2 + 2\xi\omega_2 s + \omega_2^2}{s^2 + 2\xi\omega_3 s + \omega_3^2} \right) \times \left( \frac{3.0429}{3.0430 + 3.3260 s + s^2} \right) e^{-T_{7,r} s}$	$G_0 = 6.65 \cdot 10^{-3}$ $T_{7,r} = 9.21 \cdot 10^{-1}$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$ $\omega_2 = 1.74$ $\omega_3 = 1.75$ $\xi_2 = 5.74 \cdot 10^{-3}$ $\xi_3 = 3.80 \cdot 10^{-3}$
$G_9 := \frac{G_0}{s} \left( \frac{s^2 + 2\xi\omega_0 s + \omega_0^2}{s^2 + 2\xi\omega_1 s + \omega_1^2} \right) \left( \frac{s^2 + 2\xi\omega_2 s + \omega_2^2}{s^2 + 2\xi\omega_3 s + \omega_3^2} \right) \times \left( \frac{3.0429}{3.0430 + 3.3260 s + s^2} \right) \left( \frac{1 - \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}}{1 + \frac{T_9 s}{2} + \frac{(T_9 s)^2}{12}} \right)$	$G_0 = 6.65 \cdot 10^{-3}$ $T_9 = 9.21 \cdot 10^{-1}$ $\omega_0 = 1$ $\omega_1 = 1.03$ $\xi = 1.08 \cdot 10^{-2}$ $\omega_2 = 1.74$ $\omega_3 = 1.75$ $\xi_2 = 5.74 \cdot 10^{-3}$ $\xi_3 = 3.80 \cdot 10^{-3}$

TABLE G.1 – Modélisations possibles du système mesuré

# Bibliographie

- [ACA<sup>+</sup>99] Daniel Alazard, Christelle Cumer, Pierre Apkarian, Michel Gauvrit, and Gilles Ferreres. *Robustesse et commande optimale*. Cépadués-Editions, Toulouse, France, 1999.
- [ACM84] Dennis S. Arnon, George E. Collins, and Scott McCallum. Cylindrical algebraic decomposition 1 : The basic algorithm. *SIAM Journal on Computing*, 13(4) :865–877, 1984.
- [ÅH06] Karl Johan Åström and Tore Hägglund. *Advanced PID Control*. ISA-The Instrumentation, Systems and Automation Society, 2006.
- [AHKY09] Hirokazu Anai, Shinji Hara, Masaaki Kanno, and Kazuhiro Yokoyama. Parametric polynomial spectral factorization using the sum of roots and its application to a control design problem. *Journal of Symbolic Computation*, 44(7) :703–725, 2009.
- [AHY05] Hirokazu Anai, Shinji Hara, and Kazuhiro Yokoyama. Sum of roots with positive real parts. In *Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation*, pages 21–28. ACM, 2005.
- [AL84] William F. Arnold and Alan J. Laub. Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proceedings of the IEEE*, 72(12) :1746–1754, 1984.
- [BC09] James Ward Brown and Ruel V. Churchill. *Complex Variables and Applications*. 2009.
- [BCLM14] François Boulier, Changbo Chen, François Lemaire, and Marc Moreno Maza. Real root isolation of regular chains. In *Computer Mathematics*, pages 33–48. Springer, 2014.
- [BF00] Dario Andrea Bini and Giuseppe Fiorentino. Design, analysis, and implementation of a multiprecision polynomial rootfinder. *Numerical Algorithms*, 23(2-3) :127–173, 2000.
- [BG96] Angelika Bunse-Gerstner. Computational solution of the algebraic Riccati equation. *Journal of The Society of Instrument and Control Engineers*, 35(8) :632–639, 1996.
- [BGM96] George A. Baker and Peter Graves-Morris. *Padé Approximants*, volume 59. Cambridge University Press, 1996.
- [Bha02] Grama R. Bhashyam. ANSYS Mechanical, A Powerful Nonlinear Simulation Tool. *Ansys, Inc*, 1(1) :39, 2002.
- [Bha09] Rajendra Bhatia. *Positive Definite Matrices*. Princeton university press, 2009.
- [BK13] Henri Bourlès and Godfrey K Kwan. *Linear Systems*. John Wiley & Sons, 2013.
- [Bod45] Hendrik Wade Bode. *Network Analysis and Feedback Amplifier Design*. van Nostrand, 1945.

- [BPR03] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*, volume 7. Springer Heidelberg, 2003.
- [BRB<sup>+</sup>18] Axel Barrau, Guillaume Rance, Yacine Bouzidi, Alban Quadrat, and Arnaud Quadrat. Using symbolic computation to solve algebraic Riccati equations arising in invariant filtering. *Proceedings of the 17<sup>th</sup> European Control Conference*, 2018.
- [CA76] George E. Collins and Alkiviadis G. Akritas. Polynomial real root isolation using Descartes’s rule of signs. In *Proceedings of the third ACM Symposium on Symbolic and Algebraic Computation*, pages 272–275. ACM, 1976.
- [CD82] Ming-Jeh Chen and Charles A. Desoer. Necessary and sufficient condition for robust stability of linear distributed feedback systems. *International Journal of Control*, 35(2) :255–267, 1982.
- [CD12] Frank M. Callier and Charles A. Desoer. *Linear System Theory*. Springer Science & Business Media, 2012.
- [CGY07] Jin-San Cheng, Xiao-Shan Gao, and Chee-Keng Yap. Complete numerical isolation of real zeros in zero-dimensional triangular systems. In *Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation*, pages 92–99. ACM, 2007.
- [CH91] George E. Collins and Hoon Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3) :299–328, 1991.
- [Cle15] Benoit Clement. *Commande Robuste et Contraintes d’Optimisation*. PhD thesis, Université de Bretagne Occidentale, 2015.
- [CLO06] David Cox, John Little, and Donal O’shea. *Using Algebraic Geometry*, volume 185. Springer Science & Business Media, 2006.
- [CLO15] David Cox, John Little, and Donal O’shea. *Ideals, Varieties, and Algorithms : fourth edition*, volume 3. Springer, 2015.
- [Cou97] Pierre Coustal. *Contribution au contrôle robuste des structures flexibles*. PhD thesis, 1997.
- [CZ12] Ruth Frances Curtain and Hans Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21. Springer Science & Business Media, 2012.
- [DF00] Gilles Duc and Stephane Font. *Commande  $H_\infty$  et  $\mu$ -Analyse, des Outils Pour la Robustesse*. Hermes, 2000.
- [DGKF89] John C. Doyle, Keith Glover, Pramod P. Khargonekar, and Bruce A. Francis. State-space solutions to standard  $H_2$  and  $H_\infty$  control problems. *IEEE Transactions on Automatic control*, 34(8) :831–847, 1989.
- [DLDJ<sup>+</sup>07] Michael Di Loreto, Massa Dao, Luc Jaulin, Jean-François Lafay, and Jean-Jacques Loiseau. Applied interval computation : A new approach for time-delays systems analysis. In *Applications of Time Delay Systems*, pages 175–197. Springer, 2007.
- [DS81] John Doyle and Gunter Stein. Multivariable feedback design : Concepts for a classical/modern synthesis. *IEEE Transactions on Automatic Control*, 26(1) :4–16, 1981.
- [Fau02] Jean-Charles Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, volume 2, pages 75–83, 2002.

- [FE93a] Krister Forsman and Jan Eriksson. Solving the ARE symbolically. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 363–368. IEEE, 1993.
- [FE93b] Krister Forsman and Jan Eriksson. Solving the ARE symbolically. Technical report, University of Linköping, Sweden, 08 1993.
- [FED09] Jean-Charles Faugère and Mohab Safey El Din. *De l’algèbre linéaire à la résolution des systèmes polynomiaux*. Pearson, 2009.
- [Fey13] Philippe Feyel. *Loop-Shaping Robust Control*. John Wiley & Sons, 2013.
- [Fey15] Philippe Feyel. *Optimisation des correcteurs par les métaheuristiques. Application à la stabilisation inertielle de ligne de visée*. PhD thesis, Supélec, 2015.
- [FGLM93] Jean-Charles Faugere, Patrizia Gianni, Daniel Lazard, and Teo Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *Journal of Symbolic Computation*, 16(4) :329–344, 1993.
- [FJ08] Michel Fliess and Cédric Join. Commande sans modèle et commande à modèle restreint. *e-STA Sciences et Technologies de l’Automatique*, 5(4) :1–23, 2008.
- [FÖT96] Ciprian Foias, Hitay Özbay, and Allen Tannenbaum. *Robust Control of Infinite Dimensional Systems*. 1996.
- [Fra16] Sophie Frasnedo. *Optimisation globale des lois de commande de la stabilisation inertielle d’un imageur sur critère optronique : Application à un imageur à deux étages de stabilisation*. PhD thesis, Paris Saclay, 2016.
- [FSR03] Michel Fliess and Hebertt Sira-Ramírez. An algebraic framework for linear identification. *ESAIM : Control, Optimisation and Calculus of Variations*, 9 :151–168, 2003.
- [GA94] Pascal Gahinet and Pierre Apkarian. A linear matrix inequality approach to  $h_\infty$  control. *International Journal of Robust and Nonlinear Control*, 4(4) :421–448, 1994.
- [GD88] Keith Glover and John C. Doyle. State-space formulae for all stabilizing controllers that satisfy an  $H_\infty$ -norm bound and relations to relations to risk sensitivity. *Systems & Control Letters*, 11(3) :167–172, 1988.
- [Gil91] George T. Gilbert. Positive definite matrices and Sylvester’s criterion. *The American Mathematical Monthly*, 98(1) :44–46, 1991.
- [GM88] Keith Glover and Duncan McFarlane. Robust stabilization of normalized coprime factors : an explicit  $H_\infty$  solution. In *American Control Conference*, pages 842–847. IEEE, 1988.
- [GM89] Keith Glover and Duncan McFarlane. Robust stabilization of normalized coprime factor plant descriptions with  $H_\infty$ -bounded uncertainty. *IEEE Transactions on Automatic Control*, 34(8) :821–830, 1989.
- [GR03] Alain Girard and Nicolas Roy. *Dynamique des structures industrielles*. Hermes science publ., 2003.
- [Gri15] Joseph Grifone. *Algèbre linéaire, 5<sup>ème</sup> édition*. Cépaduès, 2015.
- [GWY08] Hugues Garnier, Liuping Wang, and Peter C. Young. Direct identification of continuous-time models from sampled data : Issues, basic solutions and relevance. In *Identification of Continuous-Time Models from Sampled Data*, pages 1–29. Springer, 2008.

- [Hil08] James M. Hilker. Inertially stabilized platform technology : concepts and principles. *IEEE Control Systems*, 2008.
- [Hir13] Serge Hirwa. *Méthodes de commande avancées appliquées aux viseurs*. PhD thesis, Supélec, 2013.
- [Hol02] G. C. Holmes. The use of hyperbolic cosines in solving cubic polynomials. *The Mathematical Gazette*, 86(507) :473–477, 2002.
- [Hou59] Alston S. Householder. Dandelin, Lobacevskii, or Graeffe. *The American Mathematical Monthly*, 66(6) :464–466, 1959.
- [Ing56] Aubrey W. Ingleton. The rank of circulant matrices. *Journal of the London Mathematical Society*, 1(4) :445–460, 1956.
- [Jau01] Luc Jaulin. *Applied Interval Analysis : With Examples in Parameter and State Estimation, Robust Control and Robotics*, volume 1. Springer Science & Business Media, 2001.
- [JK85] Jan Ježek and Vladimír Kučera. Efficient algorithm for matrix spectral factorization. *Automatica*, 21(6) :663–669, 1985.
- [Kal60] Rudolf Emil Kalman. Contributions to the theory of optimal control. *Boletín de la Sociedad Matemática Mexicana*, 5(2) :102–119, 1960.
- [Kal63] Rudolf Emil Kalman. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics, Series A : Control*, 1(2) :152–192, 1963.
- [Kal64] Rudolf Emil Kalman. When is a linear control system optimal? *Journal of Basic Engineering*, 86(1) :51–60, 1964.
- [KH12] Masaaki Kanno and Shinji Hara. Symbolic-numeric hybrid optimization for plant/controller integrated design in  $H_\infty$  loop-shaping design. *Journal of Math-for-Industry*, 4(8) :135–140, 2012.
- [Kha96] Hassan K. Khalil. *Nonlinear systems*, volume 2. 1996.
- [KHO07] Masaaki Kanno, Shinji Hara, and Masahiko Onishi. Characterization of easily controllable plants based on the finite frequency phase/gain property : a magic number  $\sqrt{4 + 2\sqrt{2}}$  in  $H_\infty$  loop shaping design. In *American Control Conference, 2007*, pages 5816–5821. IEEE, 2007.
- [Kin96] Robert Bruce King. *Beyond the Quartic Equation*. Springer, 1996.
- [KS06] Masaaki Kanno and Malcolm C Smith. Validated numerical computation of the  $l_\infty$ -norm for linear dynamical systems. *Journal of Symbolic Computation*, 41(6) :697–707, 2006.
- [KYAH09] Masaaki Kanno, Kazuhiro Yokoyama, Hirokazu Anai, and Shinji Hara. Solution of algebraic Riccati equations using the sum of roots. In *Proceedings of the 2009 international symposium on Symbolic and algebraic computation*, pages 215–222. ACM, 2009.
- [Lam93] James Lam. Model reduction of delay systems using Padé approximants. *International Journal of Control*, 57(2) :377–391, 1993.
- [Lan02] Serge Lang. Algebra. *Graduate Texts in Mathematics*, 1(211), 2002.
- [Lau79] Alan Laub. A Schur method for solving algebraic Riccati equations. *IEEE Transactions on Automatic Control*, 24(6) :913–921, 1979.

- [LLMK11] Ioan Doré Landau, Rogelio Lozano, Mohammed M'Saad, and Alireza Karimi. *Adaptive Control : Algorithms, Analysis and Applications*. Springer Science & Business Media, 2011.
- [LR07] Daniel Lazard and Fabrice Rouillier. Solving parametric polynomial systems. *Journal of Symbolic Computation*, 42(6) :636–667, 2007.
- [LZ05] Ioan Doré Landau and Gianluca Zito. *Digital Control Systems : Design, Identification and Implementation*. Springer Science & Business Media, 2005.
- [Mac02] Francis Sowerby Macaulay. Some formulae in elimination. *Proceedings of the London Mathematical Society*, 1(1) :3–27, 1902.
- [Mam92] Said Mammar. *Commande multivariable robuste par les approches LQG/LTR et  $H_\infty$ . Application à un hélicoptère*. PhD thesis, Paris 11, 1992.
- [Mar77] José Martinez. Transfer functions of generalized Bessel polynomials. *IEEE Transactions on Circuits and Systems*, 24(6) :325–328, 1977.
- [MG92] Duncan McFarlane and Keith Glover. A loop-shaping design procedure using  $H_\infty$  synthesis. *IEEE Transactions on Automatic Control*, 37(6) :759–769, 1992.
- [Mig89] Maurice Mignotte. *Mathématiques pour le Calcul Formel*. Presses Universitaires de France, 1989.
- [Mor08] Guillaume Moroz. *Sur la décomposition réelle et algébrique des systèmes dépendant de paramètres*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2008.
- [Muh07] Michael Ludwig Muhler. *Robust Control System Design by Mapping Specifications into Parameter Spaces*. PhD thesis, 2007.
- [Mut80] Claude Mutafian. *Equations Algébriques et Théorie de Galois*. Librairie Vuibert, 1980.
- [Neu14] Stefan Neumark. *Solution of Cubic and Quartic Equations*. Elsevier, 2014.
- [Nic09] Richard W. D. Nickalls. The quartic equation : invariants and euler's solution revealed. *The Mathematical Gazette*, 93(526) :66–75, 2009.
- [Nou85] Jean-Pierre Nougier. *Méthodes de calcul numérique*, volume 3. Masson Paris, 1985.
- [Pad92] Henri Padé. *Sur la Représentation Approchée d'une Fonction par des Fractions Rationnelles*. Number 740. Gauthier-Villars et fils, 1892.
- [PG90] Jonathan R. Partington and Keith Glover. Robust Stabilization of Delay Systems by Approximation of Coprime Factors. *Systems & Control Letters*, 14(4) :325–331, 1990.
- [PQ04] Jean-François Pommaret and Alban Quadrat. A differential operator approach to multidimensional optimal control. *International Journal of Control*, 77(9) :821–836, 2004.
- [QQ12] Alban Quadrat and Arnaud Quadrat. Etude de l'effet du retard dans une boucle de poursuite d'un viseur gyrostabilisé. *Proceedings of CIFA2012. Grenoble*, pages 4–6, 2012.
- [Rah04] Thibaut Raharijaoana. *Commande robuste pour l'assistance au contrôle latéral d'un véhicule routier*. PhD thesis, Université Paris Sud-Paris XI, 2004.
- [RBQ<sup>+</sup>17] Guillaume Rance, Yacine Bouzidi, Alban Quadrat, Arnaud Quadrat, and Fabrice Rouillier. Explicit  $H_\infty$  controllers for 4<sup>th</sup> order single-input single-output

- systems with parameters and their application to the two mass-spring system with damping. *Proceedings of the 20<sup>th</sup> International Federation on Automatic Control World Congress*, pages 727–732, 2017.
- [RBQQ16] Guillaume Rance, Yacine Bouzidi, Alban Quadrat, and Arnaud Quadrat. A symbolic-numeric method for the parametric  $H_\infty$  loop-shaping design problem. *Proceedings of the 22<sup>nd</sup> International Symposium on Mathematical Theory of Networks and Systems*, pages 574–581, 2016.
- [RBQQ17] Guillaume Rance, Yacine Bouzidi, Alban Quadrat, and Arnaud Quadrat. Explicit  $H_\infty$  controllers for 1<sup>st</sup> to 3<sup>rd</sup> order single-input single-output systems with parameters. *Proceedings of the 20<sup>th</sup> International Federation on Automatic Control World Congress*, pages 721–726, 2017.
- [RBQQ18] Guillaume Rance, Yacine Bouzidi, Alban Quadrat, and Arnaud Quadrat. Parametric sub-optimal  $H_\infty$  controllers for an optro-mechanical system modeled by a time-delay 4<sup>th</sup> order system. *Proceedings of the 8<sup>th</sup> International Symposium on Optronics in Defence and Security*, 2018.
- [Ren89] James Renegar. On the worst-case arithmetic complexity of approximating zeros of systems of polynomials. *SIAM Journal on Computing*, 18(2) :350–370, 1989.
- [Rot09] Joseph J. Rotman. *An Introduction to Homological Algebra : second edition*. Springer, 2009.
- [Rou99] Fabrice Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Applicable Algebra in Engineering, Communication and Computing*, 9(5) :433–461, 1999.
- [Rou07] Fabrice Rouillier. *Algorithmes pour l'étude des solutions réelles des systèmes polynomiaux*. PhD thesis, Université Pierre & Marie Curie-Paris 6, 2007.
- [RZ04] Fabrice Rouillier and Paul Zimmermann. Efficient isolation of polynomial's real roots. *Journal of Computational and Applied Mathematics*, 162(1) :33–50, 2004.
- [Sch01] Carsten W Scherer. Lpv control and full block multipliers. *Automatica*, 37(3) :361–375, 2001.
- [Shm11] Sergei L. Shmakov. A universal method of solving quartic equations. *International Journal of Pure and Applied Mathematics*, 71(2) :251–259, 2011.
- [SW00] Carsten Scherer and Siep Weiland. *Linear Matrix Inequalities in Control*, volume 3. Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands, 2000.
- [Tig15] Jean-Pierre Tignol. *Galois' Theory of Algebraic Equations*. World Scientific Publishing Co Inc, 2015.
- [TVT14] Jean-Paul K. Tsasa Vangu and Yves Togba Boboy. Traitement numérique des équations de Sylvester et de Lyapunov. *LAREQ Publication*, 18 :189–199, 2014.
- [VdWAN50] Bartel Leendert Van der Waerden, Emil Artin, and Emmy Noether. *Moderne Algebra*, volume 31950. Springer, 1950.
- [Vid85] Mathukumalli Vidyasagar. *Control System Synthesis : A Factorization Approach*. Cambridge, Mass. : MIT Press, 1985.
- [Vin00] Glenn Vinnicombe. *Uncertainty and Feedback :  $H_\infty$  Loop-Shaping and the  $\nu$ -gap Metric*. World Scientific, 2000.

- [VSF82] Mathukumalli Vidyasagar, Hans Schneider, and Bruce Francis. Algebraic and topological aspects of feedback stabilization. *IEEE transactions on Automatic Control*, 27(4) :880–894, 1982.
- [WB92] Bong Wie and Dennis S. Bernstein. Benchmark problems for robust control design. *American Institute of Aeronautics and Astronautics*, 1992.
- [Zam81] George Zames. Feedback and optimal sensitivity : Model reference transformations, multiplicative seminorms, and approximate inverses. *IEEE Transactions on automatic control*, 26(2) :301–320, 1981.
- [ZDG96] Kemin Zhou, John C. Doyle, and Keith Glover. *Robust and Optimal Control*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.
- [ZF83] George Zames and Bruce Francis. Feedback, minimax sensitivity, and optimal robustness. *IEEE Transactions on Automatic Control*, 28(5) :585–601, 1983.
- [Zwi02] Daniel Zwillinger. *CRC Standard Mathematical Tables and Formulae*. CRC press, 2002.

**Titre :** Commande  $H_\infty$  paramétrique et application aux viseurs gyrostabilisés

**Mots clefs :** Commande  $H_\infty$  par loop-shaping, paramètres de modèles et paramètres inconnus, équations de Riccati algébriques, méthodes de calcul formel (systèmes polynomiaux, symbolique-numérique), systèmes à retard, viseurs, commande adaptative, géométrie algébrique réelle

**Résumé :** Cette thèse porte sur la commande  $H_\infty$  par loop-shaping pour les systèmes linéaires à temps invariant d'ordre faible avec ou sans retard et dépendant de paramètres inconnus. L'objectif est d'obtenir des correcteurs  $H_\infty$  paramétriques, c'est-à-dire dépendant explicitement des paramètres inconnus, pour application à des viseurs gyrostabilisés.

L'existence de ces paramètres inconnus ne permet plus l'utilisation des techniques numériques classiques pour la résolution du problème  $H_\infty$  par loop-shaping. Nous avons alors développé une nouvelle méthodologie permettant de traiter les systèmes linéaires de dimension finie grâce à l'utilisation de techniques modernes de calcul formel dédiées à la résolution des systèmes polynomiaux (bases de Gröbner, variétés discriminantes, etc.).

Une telle approche présente de multiples avantages: étude de sensibilités du critère  $H_\infty$  par rapport aux paramètres, identification de valeurs de paramètres singulières ou remarquables, conception de correcteurs explicites optimaux/robustes, certification numérique des calculs, etc. De plus, nous montrons que cette approche peut s'étendre à une classe de systèmes à retard.

Plus généralement, cette thèse s'appuie sur une étude symbolique des équations de Riccati algébriques. Les méthodologies génériques développées ici peuvent s'étendre à de nombreux problèmes de l'automatique, notamment la commande LQG, le filtrage de Kalman ou invariant.

**Title :** Parametric  $H_\infty$  control and its application to gyrostabilized sights

**Keywords :**  $H_\infty$  loop-shaping control, model parameters and unknown parameters, algebraic Riccati equations, symbolic computation methods (polynomial systems, symbolic-numeric), time-delay systems, sights, adaptive control, real algebraic geometry

**Abstract :** This PhD thesis deals with the  $H_\infty$  Loop-shaping design for low order linear time invariant systems depending on unknown parameters. The objective of the PhD thesis is to obtain parametric  $H_\infty$  controllers, i.e. controllers which depend explicitly on the unknown model parameters, and to apply them to the stabilization of gyrostabilized sights.

Due to the unknown parameters, no numerical algorithm can solve the robust control problem. Using modern symbolic techniques dedicated to the solving of polynomial systems (Gröbner bases, discriminant varieties, etc.), we develop a new methodology to solve this problem for finite-dimensional linear systems.

This approach shows several advantages: we can study the sensibilities of the  $H_\infty$  criterion to the parameter variations, identify singular or remarkable values of the parameters, compute controllers which depend explicitly on the parameters, certify the numerical computations, etc. Furthermore, we show that this approach can be extended to a class of linear time-delay systems.

More generally, this PhD thesis develops an algebraic approach for the study of algebraic Riccati equations. Thus, the methodology obtained can be extended to many different problems such as LQG control and Kalman or invariant filtering.